

Philips Technical Review

DEALING WITH TECHNICAL PROBLEMS
RELATING TO THE PRODUCTS, PROCESSES AND INVESTIGATIONS OF
N.V. PHILIPS' GLOEILAMPENFABRIEKEN

EDITED BY THE RESEARCH LABORATORY OF N.V. PHILIPS' GLOEILAMPENFABRIEKEN, EINDHOVEN, HOLLAND

A NEW ELECTRON MICROSCOPE WITH CONTINUOUSLY VARIABLE MAGNIFICATION

by J. B. le POOLE.

621.385.833

Several articles concerning electronic microscopes will be published in this periodical. In this first article of the series some characteristics of ordinary optical microscopes are reviewed. The wave nature of light sets a limit to the resolving power. The smallest distance between two distinguishable details is 1000 Ångström units with the best optical microscopes. A significant improvement in resolving power is obtained by using electron beams instead of light. The principle of the electron microscope is first explained with particular reference to the focusing of electron beams by magnetic lenses. A description is then given of a new electron microscope now in use at the Institute for Electron Microscopy at Delft. The advantages of the new construction over previous models are explained. The resolving power amounts to about 25 Å and the magnification is continuously variable from 1000 to 80 000 times. With this instrument it is also possible to get an electron diffraction pattern of a part of the specimen which has first been studied electron-optically, which offers the advantage of an easier identification of the materials which are being investigated. In conclusion several applications of electronic microscopes are mentioned.

In recent years investigations have been carried out in many countries with microscopes where use is made of electron beams. This has also been the case in the Netherlands, especially at the Technical University at Delft and in the Philips Laboratory at Eindhoven. It is our intention to devote several articles in this periodical to that subject. In this first article a description is given of the electron microscope that was constructed by the author for the Institute for Electron Microscopy at Delft, which is under his direction. This instrument was completed in 1944 and was then used for one month. It was then taken apart and the parts were hidden to prevent their being carried off by the enemy. Immediately after the liberation the instrument was assembled and taken into use again.

As an introduction to the description of the new microscope we shall first review several characteristics of the optical microscope for purposes of comparison, and then deal briefly with the general principles upon which the functioning of the electron microscope is based.

The optical microscope

It is a matter of general knowledge that a light microscope contains a condenser lens which con-

centrates a beam of light on the object, further an objective which forms an enlarged intermediate image of the object, and an ocular with which this image is observed. The total magnification is found by multiplying the enlargement of the objective by that of the ocular.

It is easy to see that the wave nature of light sets a limit to the resolving power, *i.e.* to the smallest distance between two details which can just be distinguished from each other. When a beam of light rays strikes an object they will be stopped by the non-transparent parts. The transparent parts transmit the rays, which, when a lens is placed in their path, converge again in the plane of the image. If, however, the opening between two non-transparent parts is of the order of magnitude of the wavelength of the light used, the light is strongly diffracted and scattered at the other side of the object over a wide angle. Because only a small part of this beam reaches the lens the part of the image corresponding to the opening will still remain dark. In this way it can be understood that the resolving power of an optical microscope is of the same order of magnitude as the wavelength of the light used. This wavelength of visible light lies between 0.8 μ and 0.4 μ .

This subject is dealt with more precisely in the theory of Abbe for the optical microscope. From this theory it follows that the resolving power

$$d_{\min} = \lambda/2A,$$

where λ is the wavelength of the light used and A is the numerical aperture of the objective.

This last quantity is equal to $n \sin u$ where n is the index of refraction of the medium in which the object lies for the light used and u is half the angle aperture of the objective.

In order to prove this we first examine how an image is formed by an objective of a grating S which is struck perpendicularly by a parallel beam of light (fig. 1). Diffraction beams will be diffracted by the object in a number of directions.

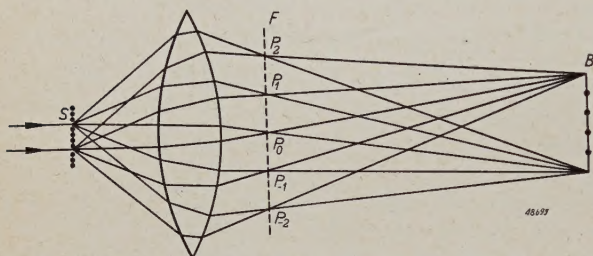


Fig. 1. Formation of the image of a grating S by an objective lens. The deflected rays converge on the focal plane F of the objective to give primary diffraction images P_0 , P_1 , P_{-1} , P_2 , P_{-2} , etc. The image B is formed by interference of the beams coming from the primary images.

Part of these beams are captured by the objective and lead to the formation of a number of diffraction images (when white light is used: diffraction spectra) in the focal plane of the objective. Abbe calls these the primary images. According to the wave theory these primary images P must be regarded as new sources of light from which issue the waves which form an image of interference in the image plane. This interference pattern is the image which is observed with the ocular of the microscope.

It is possible to calculate which diffraction beams form the primary images. Let us call the angle which a beam issuing from the grating makes with the main axis φ . According to the wave theory a large part of the light is now diffracted in very definite directions φ_k ($k = 1, 2, 3 \dots$), which are determined by the relation

$$\sin \varphi_k = \frac{k\lambda}{ny} \quad (k = 1, 2, 3, \dots),$$

where y is the linear spacing of the grating (the so-called grating constant). The first diffraction beams on either side of the main axis thus have the direction φ_1 , for which is valid $\sin \varphi_1 = \lambda/ny$.

Abbe has proved that an absolutely true image of the object is formed only when all the diffraction beams issuing from the object combine to form the image. This cannot be realized in a microscope. But a very good image is also obtained when a smaller number of beams combine.

With perpendicular illumination, in connection with the symmetrical arrangement, not only the central beam but also at least two others, thus a total of at least three beams, must

combine in the formation of the image (fig. 2a). Thus the half angle aperture u must be at least equal to the angle φ_1 . This means that $\sin u \geq \lambda/ny$ or $y \geq \lambda(n \sin u) = \lambda/A$.

This argument shows why, with perpendicular illumination, only those structures can be resolved by a microscope objective with the numerical aperture A whose mutual spacing y is at least λ/A .

Nevertheless, more can be achieved with the same objective. An impression of the grating structure can also be obtained when only two neighbouring beams cooperate, thus for example the "central" one and one of the first two diffraction beams. This can be realized by allowing the light to fall obliquely on the object in such a way that it enters the tube of the microscope just at the edge of the objective (fig. 2b). The angle φ_1 may now be twice as large as in the case just referred to. That means that in this case a grating with twice the fineness of structure can still be observed. With oblique illumination therefore the limit of the resolving power in one definite direction is given by $y = \lambda/2A$.

When a condenser is used not only a unilaterally oblique illumination is obtained, but also a universally oblique illumination. Provided the aperture of the condenser is not smaller than that of the objective, the following is then generally valid:

$$d_{\min} = \lambda/2A.$$

For an optical microscope with oil immersion A is a maximum of 1.5 for $u = 90^\circ$, so that under the most favourable circumstances one finds $d_{\min} = \lambda/3$. For such a microscope, therefore, the resolving power is about 0.2μ .

The magnification which must be used to see an object of these minimum dimensions clearly is connected with the resolving power of the human eye. If the latter is 0.1 mm, as is usually assumed, a magnification of 500 times would be sufficient. It will be preferable, however, to use a stronger

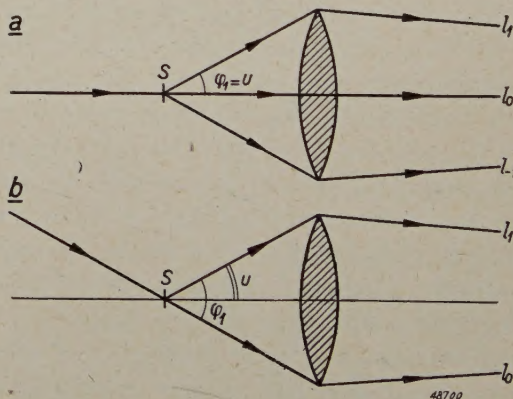


Fig. 2. The formation of primary images with a lens.

a) Upon perpendicular illumination the aperture 2μ of the objective must be at least equal to $2\varphi_1$ in order to separate the structure of the grating S .

b) With oblique illumination an image can still be formed when the aperture is equal to φ_1 . Here $\sin \varphi_1 = \lambda/ny$, where λ is the wavelength of the light used, n the index of refraction of the medium and y the grating constant. The central ray which forms the primary image P_0 (fig. 1) is indicated by l_0 ; l_1 and l_{-1} are the rays which form the primary images P_1 and P_{-1} .

magnification than strictly necessary in order not to tire the eye too much and to make every detail easily visible. A magnification of $1000 \times$ or $1500 \times$ is usually used. Greater magnifications are in general of little use with an optical microscope.

The only way of improving the resolving power when the numerical aperture is a maximum is to use light of a shorter wavelength. The use of microscopes for ultra-violet light is based on this fact. Since glass does not transmit ultra-violet rays these instruments must be equipped with quartz lenses. In this way the resolving power is increased by a factor 2.

For the distance between two details which can just be seen separately a limit of 0.1μ or 1000 \AA is then found. In many modern investigations, however, there is need of an instrument with a still higher resolving power.

Principle of the electron microscope

An important improvement in the resolving power is obtained by using electron beams instead of light.

In 1924 Louis de Broglie announced that a wave nature must be assigned to moving electrons. Experimentally this was confirmed shortly afterwards by interference experiments carried out by the American investigators Davisson and Germer. The wavelength of the electron depends upon its velocity. The greater the velocity the shorter the wavelength. According to de Broglie the wave nature of an electron is characterized by

$$\lambda = \frac{h}{mv},$$

where h is Planck's constant, m the mass and v the velocity of the electron.

When an electron with the charge e passes through a potential difference V its kinetic energy is

$$\frac{1}{2} mv^2 = eV,$$

from which it follows that

$$v = \sqrt{\frac{2eV}{m}} \quad \text{and} \quad \lambda = \sqrt{\frac{h^2}{2meV}}.$$

When we substitute the known values in this: $h = 6.6 \times 10^{-27}$; $m = 9.1 \times 10^{-28}$; $e = 4.8 \times 10^{-10}$, e.s.u., expressing V in volts, we obtain

$$\lambda = \sqrt{\frac{150}{V}} \cdot 10^{-8} \text{ cm} = \frac{12.3}{\sqrt{V}} \text{ \AA} \quad \dots \quad (1)$$

Corresponding to electrons of an energy of 150 kV, we therefore have waves of matter of about 0.03 \AA $= 3 \times 10^{-10} \text{ cm}$, thus a wavelength of the order of magnitude of hard X-rays.

Since with electron beams wavelengths are thus reached which are, for example, 100 000 times as small as those of the light used for ultra microscopy, theoretically an increase in the resolving power by the same factor is possible when the image of an object is formed with electron rays. This does not mean that this is a practical possibility. In the first place it is necessary that it should actually be possible to form an image with electron rays, and, in the second place, if that can be done, the resolving power attained still depends upon the quality of the lenses to be used for the image formation. In the case of X-rays, which also possess the very short wavelengths mentioned and therefore promise theoretically a high resolving power, the first condition is not satisfied; no medium is known which has a refractive index for X-rays appreciably different from unity. Therefore it is impossible to make "X-ray lenses". Electron rays on the other hand can be refracted and focused by means of magnetic or electrostatic fields, and thus an image can be obtained. The terminology of the optical microscope has been taken over and one speaks of magnetic and electrostatic lenses. The unavoidable errors of these lenses, as a closer theoretical consideration shows, make it impossible to obtain anywhere the limit of the resolving power that would be expected on the basis of the wavelength. With an accelerating voltage of 150 kV it will not be possible to go farther than 5 \AA . In practice, at this voltage, a resolving power of 15 to 30 \AA has already been reached. A magnetic electron lens is a short coil which thus gives a non-homogeneous magnetic field with rotational symmetry. An electrostatic lens is usually some combination or other of electrodes which gives an electrostatic non-homogeneous field, likewise with rotational symmetry. In both cases the field often has, moreover, a plane of symmetry perpendicular to the axis of symmetry. Concepts such as object distance a and image distance b , focal distance f and power $1/f$ of a lens are also defined in electron-optics in the same way as in ordinary optics. The relation between a , b and f here is also given by

$$\frac{1}{a} + \frac{1}{b} = \frac{1}{f} \dots \dots \dots (2)$$

One also speaks of thin lenses, meaning lenses so constructed that the thickness of the region within which the influence on the path of a moving electron cannot be disregarded is small compared with the object and image distance. For the power of a thin lens in the electrostatic case the following holds:

$$\frac{1}{f} = \frac{1}{8\sqrt{V_0}} \int_{-\infty}^{+\infty} \frac{1}{\sqrt{V^3}} \left(\frac{dV}{dx} \right)^2 dx \dots (3)$$

and in the magnetic case:

$$\frac{1}{f} = \frac{e}{8mV_0} \int_{-\infty}^{+\infty} H^2 dx, \dots (4)$$

where the x -axis is the axis of symmetry, V_0 is the accelerating potential, V the electrostatic potential on the axis and H the magnetic field on the axis. The magnification is given in both cases by b/a .

The fact that it is possible to focus electron rays is by no means self-evident, especially in the magnetic case where the force acting on the electron depends upon the velocity of the latter. We shall prove this below¹⁾ for the magnetic lens in a special case, namely for electrons which start from a point of the axis and follow trajectories in the neighbourhood of the axis (paraxial rays). We wish to show that all the electrons issuing from a point P_1 on the axis which lies in front of the coil, and whose trajectories make a small angle α with the axis, converge again at the other side of the lens at a point P_2 on the axis. In *fig. 3* the trajectory of one of these electrons through the field of the lens is represented diagrammatically. We call its distance from the axis r and its velocity v . The small radial velocity of the electron then amounts initially to va . The distances from P_1 and P_2 to the middle of the lens are respectively a and b . From the fact that we confine ourselves to paraxial rays it also follows that the maximum distance r_0 to the axis is so small that at that distance the axial component H of the magnetic field is practically the same as on the axis.

As soon as the electron enters the magnetic field it is deflected laterally. It will then travel along a helix around the axis with a gradually changing radius. The change in the radius, which amounts to a movement of the electron towards the axis, is caused by a force which is the result of the axial magnetic field and the angular velocity with respect to the axis. This angular velocity $\dot{\varphi}$ can be calculated in the following way. The change per unit of time in the impulse moment of

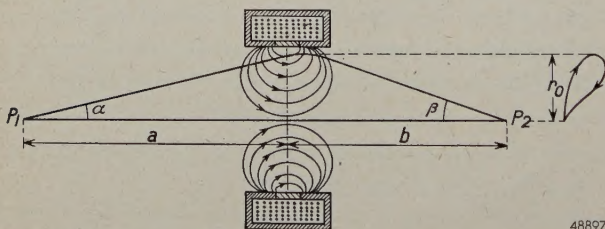


Fig. 3. Diagram of the path of an electron in a thin magnetic lens: projection on a plane through the axis of symmetry (left) and projection on a plane perpendicular to this axis (right). The lines of force drawn characterize the magnetic field of the lens. The electron leaves the axis at P_1 and reaches it again at P_2 . The distances from P_1 and P_2 to the middle plane of the lens are a and b , respectively; the angles which the ray make with the axis at P_1 and P_2 are α and β respectively. The maximum distance of the electron from the axis is r_0 .

¹⁾ A similar proof may be found also in A. Bouwers, *Physica* 4, 200, 1937.

the electron with respect to the axis is equal to the moment of the Lorentz force with respect to the axis. Hence

$$\frac{d}{dt} m r^2 \dot{\varphi} = r (e v_r H - e v H_r), \dots (5)$$

where v_r is the radial velocity and H_r the radial magnetic field (calculated as positive in the direction of the axis). Here $v \cos \alpha$ is replaced by v .

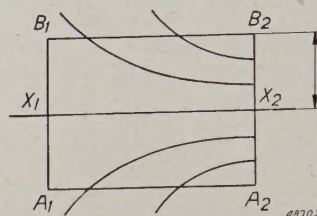


Fig. 4. The magnetic lines of force in a small cylinder whose axis coincides with the axis of symmetry of the magnetic field. The cylinder with the radius r is shown in cross section. A_1B_1 and A_2B_2 represent two circular cross sections of the cylinder perpendicular to its axis at the points x_1 and x_2 . The axial components of the field strength at x_1 and x_2 are H_1 and H_2 respectively.

On the basis of *fig. 4* we shall now calculate this H_r . In the figure A_1B_1 and A_2B_2 represent two circular cross sections with the radius r taken perpendicular to the axis of the coil at the points x_1 and x_2 . The axial components of the field strength at x_1 and x_2 we shall call H_1 and H_2 respectively. The number of magnetic lines of force passing through the cross section A_2B_2 but not through A_1B_1 is $\pi r^2 (H_2 - H_1)$. This is the number of lines of force cutting the cylindrical surface with the area $2\pi r(x_2 - x_1)$. The average value of the radial magnetic force is therefore $H_r = r(H_2 - H_1) / 2(x_2 - x_1)$, from which, upon passing over to an infinitesimal distance, one finds that

$$H_r = -\frac{r}{2} \cdot \frac{dH}{dx} \dots (6)$$

The above equation of motion (5) thus becomes

$$\frac{d}{dt} r^2 \dot{\varphi} = \frac{e}{m} \left(r \frac{dr}{dt} H + \frac{r^2}{2} \frac{dH}{dx} \frac{dx}{dt} \right).$$

The expression between parentheses in the second member is nothing else than the differential quotient of the product $\frac{1}{2} r^2 H$ with respect to time, so that we obtain

$$\frac{d}{dt} r^2 \dot{\varphi} = \frac{e}{m} \frac{d}{dt} \frac{1}{2} r^2 H.$$

From this it follows that

$$\dot{\varphi} = \frac{e}{2m} H \dots (7)$$

We see, therefore, that the angular velocity of the electron is proportional to the axial magnetic field. It becomes zero as soon as the electron has passed through the field of the coil. The electron cannot pass the axis. Behind the coil it moves in a plane which passes through the axis and then, when the field of the coil is strong enough, it must strike the axis due to its inwardly directed velocity.

In order to find the inwardly directed acceleration of the electron it must be noted that a centripetal force $m\dot{\varphi}r^2$ would be necessary to keep the electron at a constant distance r from the axis. But the axial magnetic field and the lateral velocity together cause an inwardly directed Lorentz force:

$$-K_r = e \dot{\varphi} r H.$$

The resulting acceleration is therefore

$$\ddot{r} = K_1/m + \dot{\varphi}^2 r,$$

or, according to (7)

$$\ddot{r} = -\frac{e^2 r}{4m^2} H^2 \dots \dots \dots (8)$$

For the case of a thin lens, thus where the field acts only over a short distance, r may be considered a constant r_0 . Equation (8) can then easily be integrated. By setting $dt = dx/v$ — which is permissible, because we have confined ourselves to rays in the neighbourhood of the axis — we find that

$$\dot{r} = \frac{v r_0}{a} - \frac{e^2 r_0}{4m^2 v} \int_{-\infty}^{+\infty} H^2 dx \dots \dots \dots (9)$$

Due to this velocity the electron will reach the axis after the time $\tau = r_0/\dot{r}$. Since it is found from formula (9) that this time is the same for all electrons issuing from a point P_1 on the axis, they are indeed all focused at the same point P_2 , which concludes the proof.

For a thin lens it is now easy to prove formulae (4) and (2). Just as the electron has the radial velocity $\dot{r} = vr_0/a$ before entering the magnetic field, thus between P_1 and the lens, after leaving the magnetic field, thus between the lens and P_2 , it will have the radial velocity

$$\dot{r} = -\frac{v r_0}{b},$$

b being the image distance, i.e. the distance between P_2 and the lens. When this is substituted in (9) we find that

$$\frac{1}{a} + \frac{1}{b} = \frac{e^2}{4m^2 v^2} \int_{-\infty}^{+\infty} H^2 dx,$$

which corresponds to (2) when the expression (4) is substituted for $1/f$ and $\frac{1}{2}mv^2 = eV_0$ is taken into account.

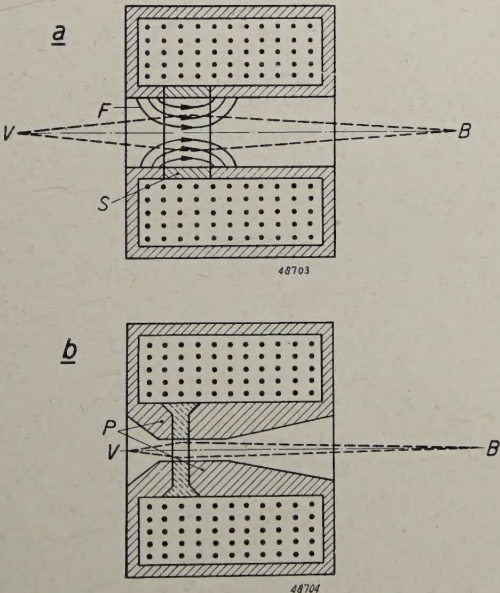


Fig. 5. Drawing of a cross section of a magnetic lens enclosed in an iron jacket, *a* without, and *b* with pole shoes. In these figures F are the lines of force, S is the slit, P the pole shoes. At V is the object and at B the image. In case *a* the focal distance may amount to several centimetres and in *b* to several millimeters. In connection with the vacuum which is maintained inside the microscope, the slit in the iron jacket is closed by a copper ring.

In order to obtain strong magnetic lenses the coil was encased in an iron jacket, just leaving a narrow annular slit on the inside. *Fig. 5a* is a diagram of this construction. The whole magnetic field is then concentrated on the immediate vicinity of the slit. At the acceleration voltages customary in

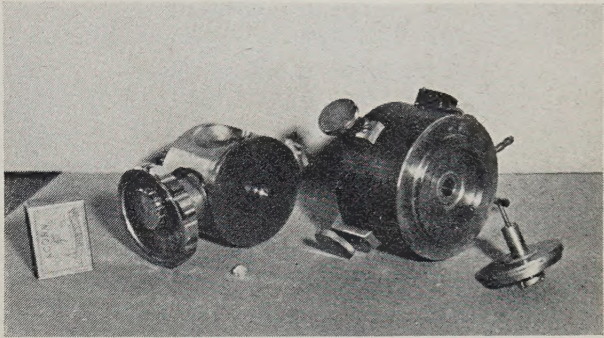


Fig. 6. Parts of a modern electron microscope. On the left a sample holder with air lock. The substance is placed on the point at the middle of the front surface of the holder. On the right an objective, and lying in front of it the holder for the objective diaphragm.

electron microscopy such iron-clad coils can be used as magnetic lenses with a focal distance of several centimeters. In order to increase the power of such a lens still more pole shoes are introduced at the slit (fig. 5*b*). In this way the focal distance is reduced to not more than a few millimeters. It is easy to understand that with the construction sketched in fig. 5 the total number of ampere turns is a measure of the strength of the field between the pole shoes. By varying the current the strength of the magnetic field is changed and thereby also the focal distance of the lens. A magnetic lens of modern construction is shown in fig. 6.

The arrangement of an electron microscope

The first electron microscopes were constructed in Germany around 1934. Later on this work was continued in other countries. *Fig. 7* gives a diagrammatic representation of the arrangement of such an instrument with magnetic lenses.

The source of electrons is usually so arranged that the electrons emerging from the cathode get an energy of 50 to 100 kV. They pass through the circular opening in the anode and leave the accelerating tube as a beam with a constant velocity. The magnetic field of a condenser lens concentrates the electrons more or less, as desired, on the object situated close above the objective. In order to obtain a sufficiently narrow beam — the marginal rays blur the image — a physical aperture is usually placed in the condenser.

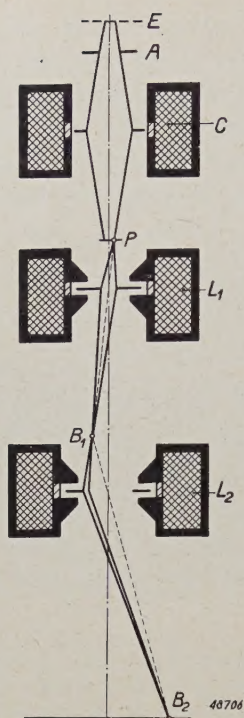


Fig. 7. The path of the rays of a magnetic electron microscope. *E* is the source of electrons, *A* the anode, *C* the condenser, *L*₁ the objective, *L*₂ the projector, *P* the object, *B*₁ the intermediate image, *B*₂ the final image.

The electrons, which converge on the object within a small area, are scattered by the different parts of the object in varying degrees. In most instruments the rays which are scattered too much in a part of the object strike a diaphragm and thus do not contribute to the formation of the image. At the place in the image corresponding to that part of the object a dark spot is then left. The other electrons which upon passing through the specimen are scattered only over small angles then pass through the objective and are focused into an image by the magnetic field of that lens. It can be shown that the solid angle over which the electrons are scattered is proportional to the mass through which the rays have passed. The contrast effect in the image is thus based upon a difference in mass in the parts of the specimen.

In most microscopes this intermediate image formed by the objective is made visible on a fluorescent screen situated at *B*₁ in fig. 7. There is then a side opening through which it can be observed visually. For this first stage a magnification of 100 to 125 \times is usually employed.

The above-mentioned screen is situated immediately above a second microscope lens (disregarding the condenser), which is called the projector, and has a hole in the centre. The part of the image lying above this hole passes through the projector and is further magnified on the fluorescent screen

at the bottom of the instrument. This final image can be observed through a side tube likewise provided with a glass window. The magnification of the projector, which is only slightly variable in each instrument, lies between 50 and 300 \times as needed. It is obvious that if desired the last mentioned fluorescent screen can be replaced by a photographic plate.

The whole tube in which the electrons move must be evacuated as completely as possible, since otherwise the electrons would be very much disturbed in their motion. During the observation the pressure must remain below 0.001 mm of mercury. With the help of a so-called air lock a new object can be put into position without seriously disturbing the vacuum.

It is desirable to take measures for the stabilisation of the high tension. Since the focal distance depends upon the acceleration voltage (cf. formula (4)), the image would otherwise be blurred owing to fluctuations of the voltage.

From the same formula it follows that the strength of a magnetic lens also depends upon the field strength. From that it follows that it is essential to keep the current in the lenses highly constant, because otherwise their focal distance would vary. This can be achieved by taking care that the lens current is supplied by a source with constant voltage, while water-cooling prevents the temperature and thus the resistance of the coils from changing too much.

Some investigators have preferred electrostatic lenses. The choice between the two systems is determined by the requirements made of the instrument. When a high magnification is desired magnetic lenses offer some advantage over electrostatic ones. Magnetic lenses have been successfully constructed with a very short focal distance, whereas until now this has proved very difficult with electrostatic lenses. With magnetic lenses the spherical aberration is also less. We shall return to this later.

Further, the strength of magnetic lenses, in contrast to that of electrostatic lenses, can easily be varied from the outside, which is an advantage for adjusting magnification and focusing.

On the other hand, an advantage of electrostatic lenses is that the supply voltage of the accelerating tube and the lenses need not be so rigorously stabilized²⁾.

²⁾ From the above formula (3) it would even follow that if V_0 and V come from the same source (mains voltage) no stabilization is necessary at all. Secondary effects, among which a correction to be introduced on the basis of the theory of relativity, make it necessary, however, that there should still be stabilization (to 0.5%).

The instrument described in this article is equipped with magnetic lenses.

The use of an intermediate lens

In the discussion of the new electron microscope we shall give special attention to the properties which distinguish it from the electron microscopes so far described in literature.

An important improvement is the introduction of a third lens between objective and projector which makes it possible to vary the magnification continuously between wide limits.

In order to make this clear we shall first explain that in a two-stage microscope it is impossible to alter the magnification over a wide range. This is connected with the aberrations occurring under certain circumstances and with the construction of the microscope.

In electron microscopes aberrations occur which are analogous to those in optical lenses. One of the most important of these is the spherical aberration, resulting from the fact that the edge zones of a lens are stronger than the centre. By increasing the field strength of a lens, its focal distance becomes smaller, and with the same angular aperture a part of the magnetic field lying closer to the axis is used. The influence of spherical aberration then becomes less. It is therefore desirable to increase the number of ampere turns as much as possible. A limit is set to this by the magnetic saturation taking place in the pole pieces. Now in order to avoid this, all the dimensions of the lens could be increased, but this raises the difficulties that the lenses then consume too much energy and the microscope assumes too large proportions. A compromise must therefore be sought. A satisfactory solution can be found, at the voltage of 150 kV chosen for our instrument, by using lenses with about 3000 ampere turns. With such lenses a focal distance for the objective of $f = 4.5$ mm has been obtained. For the other lenses a focal distance of 3.5 mm can be attained, because in those cases there are no difficulties connected with the introduction of the specimen.

From calculations of Glaser and Dosse it is found that only about $1/4$ to $1/3$ of the inner diameter of a magnetic lens may be used if it is desired to prevent distortion, an image aberration closely connected with spherical aberration, from being larger than 5 to 10 percent, a limit which is considered permissible also in the optical microscope. As a result of this limitation the magnification of a magnetic lens in a microscope can be regulated within rather narrow limits, as we shall prove below.

In the case of the projector of the microscope described here the diameter of the bore amounts to 3.5 mm, so that the diameter of the useful field, *i.e.* of the cross section of the beam permissible within the lens, is 1 mm. The diameter of the final screen amounts to 180 mm. If this is to be used to its full extent, therefore, a magnification of $180\times$ is necessary. The projector has therefore been constructed for this normal maximum magnification; with a further increase in power the strength of the projector increases only slightly. It is of course possible to obtain an image with a lower magnification, but then only part of the final screen is filled, so that the possibilities of the instrument are not fully utilised.

The magnification by the objective is entirely fixed. In this case the position of the object is fixed, while the image distance is determined by the position of the object plane of the projector lying immediately above that lens.

In a two-stage microscope the magnification can therefore only be varied within narrow limits.

When it is required of a microscope that the magnification shall be completely controllable, special measures must be taken. It is advisable to find a solution where the projector, which in any case can only be changed to a slight extent, is maintained at a perfectly constant strength. When this is done it is possible to adapt this lens entirely to the magnification it is required to give, *i.e.* to construct it in such a way that with respect to distortion it is corrected to a high degree, while the largest possible field angle is used.

A simple method of varying the magnification is now obtained by adding an extra lens between objective and projector. When the projector is kept constant the aperture of the projector may be considered as the projection screen of the "two-stage microscope" consisting of the objective and the new lens. Because the projector has a high magnification this ocular aperture will be small, even when a reasonably large final screen is provided. We have already stated that in our case the diameter of the bore of the projector amounts to 3.5 mm and consequently that of the ocular aperture amounts to about 1 mm.

The extra lens — which we shall call the intermediate lens — need only fill this small field free of distortion. In the case of the two-stage microscope discussed above (formed by objective and projector) the possibility of regulation was only slight because of the requirement made that the final screen of 180 mm diameter should always be filled. Without that requirement variation of the

magnification would also have been possible with that microscope. With the two-stage microscope formed by objective and intermediate lens the requirement is now made that it shall merely fill a "screen" of 1 mm diameter. It will be clear that this can easily be satisfied also with a great variation of magnification.

The introduction of an intermediate lens not only makes it possible to vary the magnification, but also the maximum magnification now considerably exceeds that which is attainable with two lenses. In the case of the microscope described here, objective and projector together give a magnification of 6000 times. By changing the powering of the intermediate lens the magnification can be varied from 6000 to 80 000 times. This total magnification is of course the product of the magnifications of the three separate lenses.

The trajectories when three lenses are used are shown diagrammatically in *fig. 8a*.

Is it now also possible by means of the intermediate lens to vary the magnification from 6000 times to smaller values? Theoretically it certainly is. When the power supplied to the intermediate

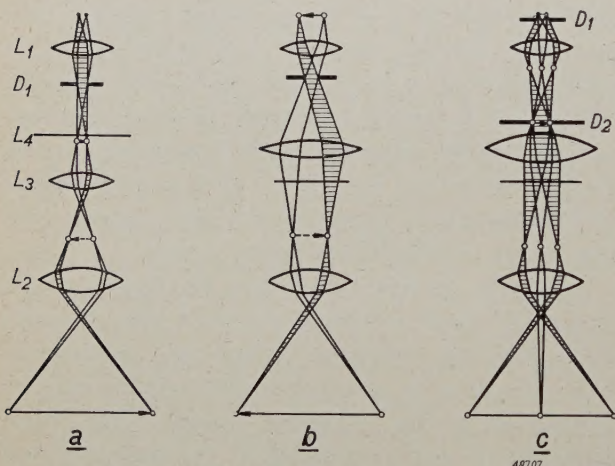


Fig. 8. Diagram of the path of the rays in the electron microscope described in this article.

L_1 is the objective, L_2 the projector, L_3 the intermediate lens, L_4 the diffraction lens, D_1 the objective diaphragm and D_2 the diffraction diaphragm.

a) The intermediate lens is in action. This lens converges the rays so strongly that they meet above the projector and then form a new intermediate image of the image formed by the objective lens. The intermediate lens makes it possible to vary the magnification continuously between $6000\times$ and $80\,000\times$.

b) The diffraction lens deflects the rays which have passed through the objective, before they have formed an image, and thereby reduces the size of the image. The magnification is variable from $6000\times$ to $1000\times$.

c) The microscope is used for studying a diffraction pattern. As for image formation, the setting is for the magnification 0. The diffraction image, which is formed as "primary image" in the focal plane of the objective, is projected in magnified form on the final screen by the diffraction lens and the projector. The objective diaphragm is now pushed up to allow the passage of the diffraction beams.

lens is reduced its magnification decreases. The point of intersection of the rays, seen in *fig. 8a*, between the intermediate lens and the projector then shifts towards the projector. When the power of the intermediate lens is further reduced there comes a moment when the above-mentioned point of intersection of the rays lies in the object plane of the projector. This position is very important, because, as will appear in the continuation of this article, a diffraction image is then obtained. The magnification is then zero. As the power is further reduced the point of intersection falls still lower. Upon passing through the magnification zero the image is inverted. The magnification then increases again until finally no power current at all flows through the intermediate lens and we have once more the two-stage microscope with which we began. In practice, however, there is a serious objection to this method. When the point of convergence of the rays is shifted to a lower level, an increasingly large part of the bore of the intermediate lens has to be used. (The final screen must in any case be filled.) Therefore the distortion increases. At low magnification, therefore, this image aberration becomes much too large. This could be remedied by making the bore of the pole pieces of the intermediate lens larger. Then, however, its minimum focal distance also becomes larger, so that the maximum magnification of this lens is reduced.

A better solution is to use another intermediate lens with a large bore for the range of low magnifications. For these reasons a fourth lens was introduced into the microscope. Since this is also used for making diffraction photographs (this will be discussed later) it is called the diffraction lens. It is situated between the objective and the intermediate lens. When this fourth lens is used the intermediate lens is switched off. By varying the current in the diffraction lens the magnification can be made variable for all values between 6000 and 1000 times. The magnification of the instrument is thus as a whole continuously variable from 1000 to 80 000 times. The lower limit gives the desired transition to the magnification of the optical microscope.

In *fig. 8b* the path of the rays is shown diagrammatically for the case where the diffraction lens is used. This lens is given such a power that the point of convergence of the rays mentioned above lies below the object plane of the projector. The diffraction lens is weaker than the intermediate lens. It does not give a real intermediate image. It refracts more or less the rays coming from the objective and in this way reduces the

magnification as compared with that of the original two-stage microscope.

We shall now first discuss the second function of the diffraction lens.

The recording of diffraction patterns

In 1927 G. P. Thomson demonstrated in a perfect manner the correspondence between moving electrons and short electromagnetic waves (X-rays), when he showed that it is also possible to obtain the well-known Debye-Scherrer rings when beams of fast electrons are sent through thin foils.

The significance of these diffraction diagrams for the identification of substances and the study of their crystalline state has already been discussed in this periodical³⁾.

When a sample of a substance is investigated with an electron microscope it is in many cases desirable to have a diffraction pattern of the substance at one's disposal at the same time.

When a diffraction diagram is to be made with the electronic microscopes so far described in literature the lenses are switched off. But the small diameter of the pole pieces of the projector makes it impossible to record a diffraction pattern immediately, because the diffraction rings are then cut off. Thus either these pole pieces must be removed, which takes some time, or the sample must be placed in an air lock anew under the projector, or an extra plate camera has to be introduced above the projector, which makes the construction much more complicated. Only in the last case is it still

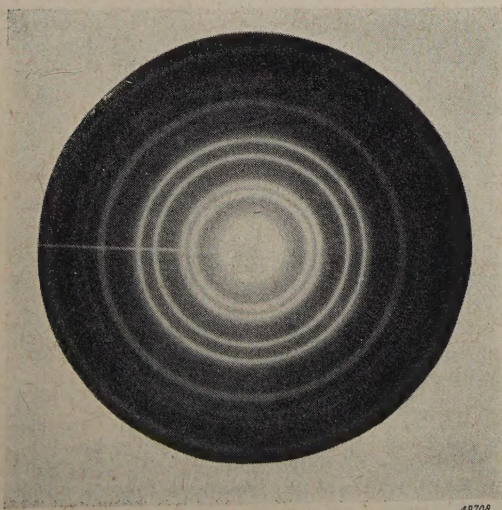


Fig. 9. Diffraction diagram of a gold sol taken with the instrument described in this article.

³⁾ W. G. Burgers, X-rays and electron rays as aids in chemical and metallographic investigation. Philips Techn. Rev. 5, 161, 1940.

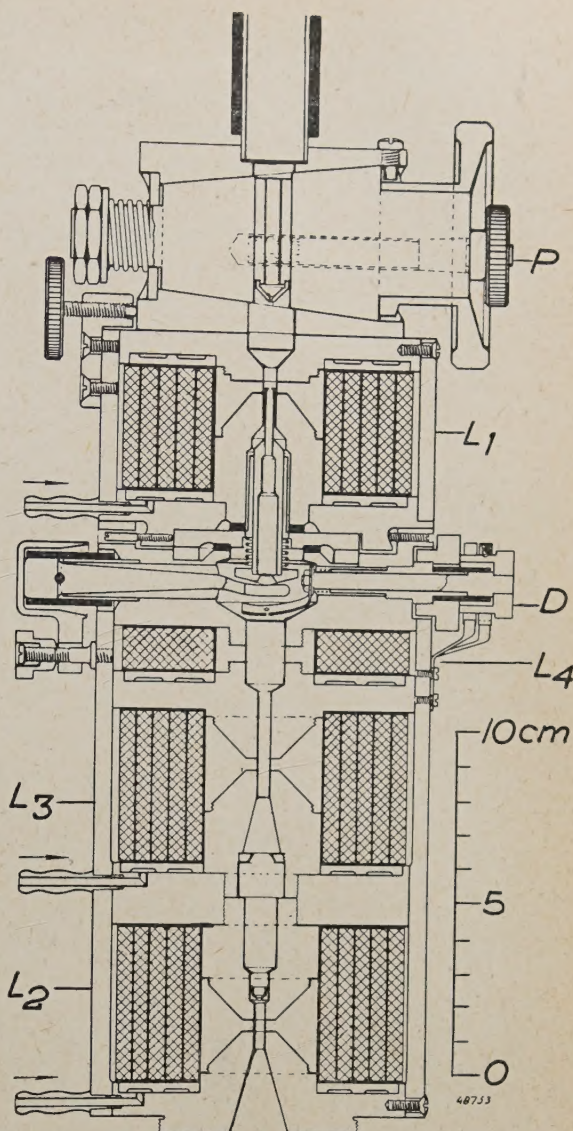


Fig. 10. Drawing of a cross section of the microscope. The magnetic lenses may be seen whose action is described in the text. *P* is the lock for the object, *L*₁ the objective, *L*₂ the projector, *L*₃ the intermediate lens, *L*₄ the diffraction lens, *D* the diffraction diaphragm. Since the magnification is easily variable with the help of intermediate and diffraction lenses, no use is made of the intermediate image formed between objective and projector for seeking the detail of the object to be magnified. This intermediate image is not, therefore, made visible on a fluorescent screen, as is customary in most other constructions.

possible to find the diffraction image of a part of the specimen of which the electron-optically magnified image has first been observed.

The introduction of an extra lens as described above makes it possible to simplify considerably the method of taking diffraction diagrams. We then make use of the weak lens called the diffraction lens. With its help the focal plane of the objective is projected on to the object plane of the projector (fig. 8c). In this focal plane, as follows from the theory of Abbe, lies the primary image and, with sufficiently wide aperture, therefore also the

diffraction image (*cf.* fig. 1). This diffraction image is very small owing to the small focal distance of the objective, but the projector enlarges it to the size customary for such patterns. *Fig. 9* gives an example of a diffraction pattern obtained in this way.

In order to obtain good image formation with sufficient contrast when using the instrument in a normal way, it is necessary to screen off the rays that are too strongly scattered by the specimen. For that purpose a diaphragm with an aperture of about $70\ \mu$ is introduced in the objective, which

can be accurately adjusted from the outside. But when diffraction is employed it is just the strongly scattered rays which are needed. For that purpose the objective diaphragm is raised and thus put out of action, the diaphragm in the diffraction lens then being brought into play. In this lens there are two diffraction apertures of different size, which can be used at discretion. When the small aperture is used it is possible to make diffraction diaphragms of very small regions, for when the object is sharply focussed on this diaphragm only those electrons can pass through which come from the corresponding part of the object. In the case of the microscope described here the magnification on the diaphragm is $13\times$. When the aperture of $40\ \mu$ is used it is possible to scan the object and to cut out regions with a diameter of $3\ \mu$.

The great advantage of this construction is that it is easy to switch over from the electron image to the diffraction image and that one can then observe a diffraction pattern of that part of the object whose image has first been studied microscopically. This furnishes a good method for the identification of the crystals in the specimen being investigated.

Some further details of the construction

Fig. 10 shows a cross section of the microscope and *fig. 11* is a photograph of the instrument.

For the maximum accelerating voltage $150\ \text{kV}$ was chosen because this gives the electrons a reasonably large power of penetration and the necessary high-tension installation had already been developed by the Philips Physical Laboratory⁴⁾.

The current for the lenses is supplied by an accumulator battery of 35 volts with a sufficiently

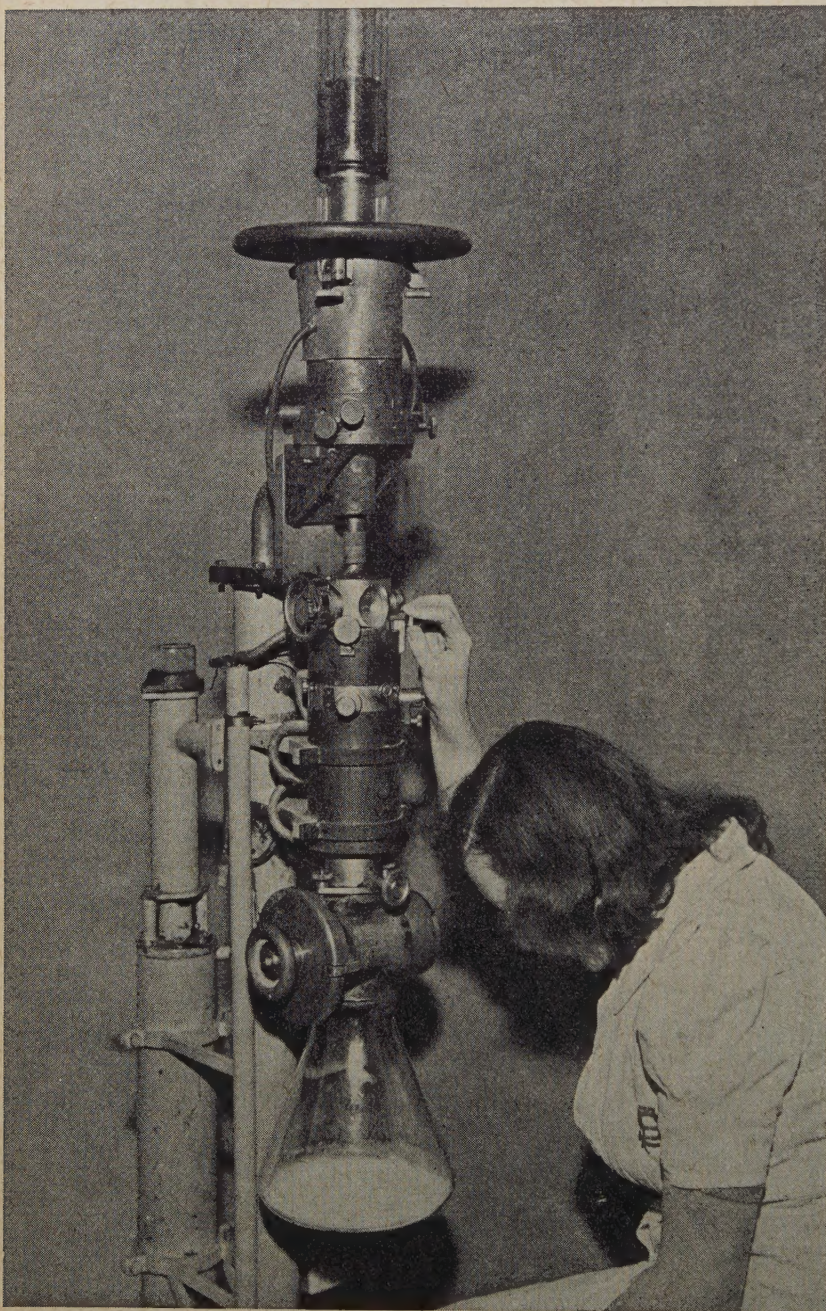


Fig. 11. The electronic microscope for $150\ \text{kV}$ in the Institute for Electronic Microscopy at Delft.

⁴⁾ This installation, which furnishes a voltage constant to within 0.2% , will be described by A. C. van Dorsten in one of the following numbers of this periodical.

constant voltage. The maximum current per lens is about 1.4 ampere. Each of the lenses is cooled separately. In order to obtain a rapid dissipation of heat, layers of copper are introduced between the windings at regular intervals. The conduction of heat towards the cooled side walls is thus increased, so that a current density of 6.5 A/mm^2 became permissible.

The introduction of a third lens between objective and projector also gives the advantage that the length of the microscope tube could be considerably decreased. The total distance from the object to the final image in the new microscope amounts to 600 mm, which is short considering that the diameter of the final image is 180 mm.

The electrons thus have short trajectories, and that also has several advantages. In the first place there is less chance of collision with residual gas molecules in the tube, and in the second place the influence of disturbing fields is less. The path of the rays between the objective and the next lens is most sensitive to these fields. Its length in our microscope amounts to only 10 cm, compared with 30 cm in the model which until now was the best in that respect. Moreover, owing to the compact structure the path of the rays is doubly shielded by the tube itself and by the coil jacket. As a third advantage of a short microscope tube the small volume may be mentioned, which makes it possible to reach a high vacuum more quickly.

The coils of the various lenses are separated by the iron side walls of the jackets. Only between the objective and the diffraction lens is a small space left free in which several essential components are housed: in the first place the three adjusting screws by means of which the objective can be aligned with respect to the other lenses; further a similar arrangement for adjusting the objective diaphragm — here is also located the fork with which the last mentioned diaphragm can be moved up for making diffraction photographs — and finally in this space the diaphragm holder is housed with the two already mentioned apertures of different size for the electron diffraction. The larger or the smaller of these can be used at will, or the passage can be made entirely free. All these arrangements can be operated while the microscope is in use.

In electron microscopy the visual image is usually only used to obtain an impression of the specimen and for sharp focusing. The most important observations, however, are made photographically. A photographic plate or film must then be brought into the path of the rays.

It is desirable to obtain sharp photographs of

the whole image. As already stated, the diameter of the final image is 180 mm. It would be difficult to introduce a photographic plate of that size into the vacuum close to the final screen. We therefore use a 35 mm film, which, in order to be able to cover the whole image, is introduced at a spot where the cross section of the beam is still sufficiently small. Because of the small apertures customary in electron microscopy, the depth of focus is more than sufficient to obtain a sharp image at that spot without a new adjustment. In this way the exposures are made with $1/5$ of the total magnification.

This must not, however, by any means be considered as a disadvantage. The resolving power of the film is much better than that of the eye. It is therefore possible subsequently to enlarge the small photographs more than five times.

The fact that the camera is placed close to the projector has also the advantage that because of the high current density of the beam impinging on the film a short exposure is sufficient.

It is obvious that in visual observations the camera must be removed from the beam of the rays. It was found possible to introduce a simple arrangement with which the film holder can be tipped

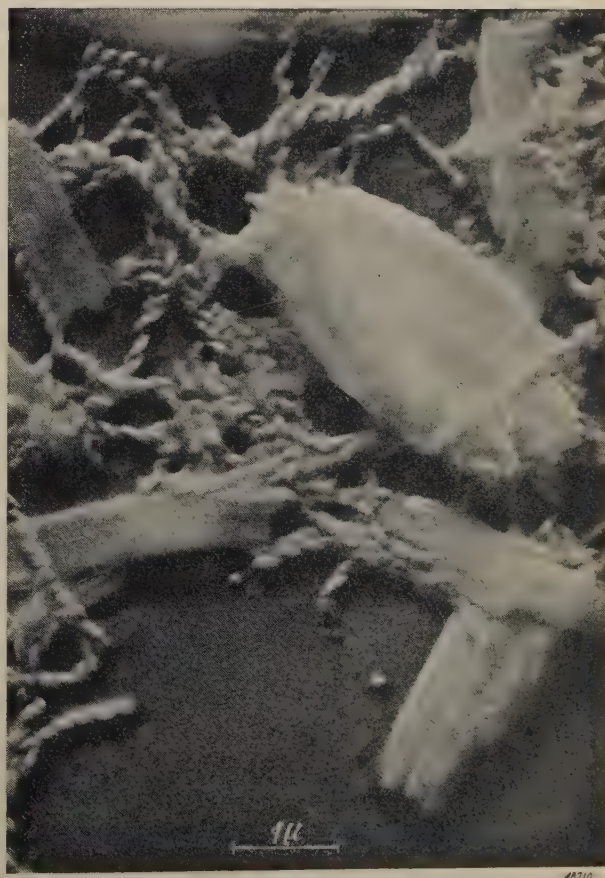


Fig. 12. A sample of fat magnified 20 000 times electron-optically with an accelerating voltage of 90 kV.

upwards from the outside. In order to prevent double exposure the film is automatically shifted when the camera is tipped up. When 25 photographs have been made a new film must be placed in the vacuum. Then the whole microscope has to be exhausted anew. Although in this process the gases which the new film brings with it have also to be removed, this operation does not take longer than about 10 min.

One of the difficulties attending the use of high magnifications is the low brightness of the final image. This makes the focusing very difficult. A special focusing arrangement has therefore been introduced.

By means of the electric field between two sets of parallel plates situated between the condenser and the objective lens, the incident ray is caused to oscillate back and forth with a frequency of 50 c/sec. When the microscope is not exactly focused this oscillation will blur the image. The power current of the objective is now varied until, in spite of the oscillation of the incident ray, the image remains sharp, whereupon the voltage on the deflection plates is switched off again.

The focusing arrangement is so constructed that the incident beam oscillates over an angle of $1/100$ radian. Very good results are obtained with this method. Since in this way rays which have an aperture 20 times as large as the beam used for image formation are focused as well as possible, the adjustment can be improved by a factor of 20.

In conclusion it may be mentioned that the resolving power of the microscope lies between 25 and 30 Ångström. It will be possible to approach the theoretical limit of 5 Å, mentioned at the beginning of this article, more closely by improving the rotational symmetry of the lenses.

In *figs. 12 and 13* two photographs are reproduced which were taken with this instrument.

Summarizing, the most important advantages of the electron microscope described here as compared with previous types are the following:

- 1) The magnification can easily be continuously varied between $1000\times$ and $80\,000\times$.
- 2) The length is shorter and therefore the effect of disturbing fields is also smaller.
- 3) The image field is much larger.
- 4) With the 35 mm camera it is possible to make a large number of exposures in a short time.
- 5) A special arrangement provides for accurate focusing.
- 6) It is possible to pass over immediately from electron to diffraction pattern and to make a

diffraction pattern exposure of a selected region of $3\,\mu$ diameter in the sample whose image has first been studied microscopically.

Applications of the electron microscope

The electron microscope has opened up entirely new fields of investigation for various branches of science. Its applications are widely varied and we can only mention a few of them.



Fig. 13. Molybdenum oxide magnified 30 000 times electron-optically with an accelerating voltage of 82 kV.

Although the idea of examining specimens with such a high magnification is very attractive, some investigators at first had objections to this new instrument. They feared that it would make a difference having to observe specimens in a vacuum. In order to settle this, various kinds of specimens were also investigated with an optical microscope in a vacuum. It was found that in most cases the vacuum presents no difficulty. Primitive organisms such as bacteria remain alive in spite of a certain drying out. It was also asked whether, owing to the continuous electron bombardment, the specimens might not be fundamentally changed. It is possible to determine whether changes occur during an observation. When sufficiently thin preparations are examined and the radiation intensity is limited, there are usually no serious objections. The bacteria will indeed be killed by the electron bombardment, but with the optical microscope the bacteria are usually dyed and this likewise kills them.

In medical science it was long a difficulty that various diseases are caused by organisms or

substances which are invisible under an ordinary microscope. Such a virus can as a rule be observed with an electron microscope.

With this instrument it is also possible to observe structural elements only a few $m\mu$ in size in tuberculus bacillae and other bacteria, which provides the possibility of distinguishing morphologically different types of disease producers. Further, medical investigators and biologists have made use of the instrument to study the structure of muscles and nerve fibrilla as well as the structure of protoplasm and the problems of chromosome structure and the arrangement of the genes. It was also very important that various not yet completely explained processes, such as the occurrence of coagulation in blood, could be studied under very high magnification.

Many plant diseases are also caused by viruses. The virus of the tobacco mosaic disease and that of a very destructive potato disease have, among others, been made visible with the electron microscope. This has led to great advances being made in phytopathology in recent years. The viruses often prove to be rodlike in shape and of the order of magnitude of molecules.

It has also been possible with an electron microscope to make various bacteriophages visible. These are kinds of viruses which destroy certain bacteria. This process — called bacteriophage lysis — has also been investigated with the electron microscope.

Bacteria had already long been observed with the optical microscope, but the cilia of these organisms, which have a thickness of from 0.02 to 0.02 μ , and all kinds of details and complicated structures of the bacterial cell were

only made clearly visible with electron rays.

For mineralogy and technology the new microscope is also an important aid. In the problems there encountered it is often desired also to be able to observe a diffraction diagram in order to identify the crystals with the help of the Debije-Scherrer rings. It is of great advantage when one can easily switch over from the electron to the diffraction image.

In the field of silicates, by the application of the new method discoveries have been made which are of great importance in ceramics.

It is important to note that here there has been a transition from qualitative to quantitative methods. The adsorption power of the surface of particles of clay has often to be investigated. One then counts and measures under an electron microscope, for instance, 1000 crystals of a substance and makes a distribution curve of their sizes. In making such distribution curves it is very desirable to be able to change over easily to a lower magnification. The study of chemical processes at a magnification of 20 000 to 30 000 \times is also very instructive. As an example we may mention an excellent investigation of the manner in which an image is formed on a photographic plate, which was carried out by Von Ardenne before 1940.

It may be assumed that within a reasonable time every large hospital and every laboratory for research in the field of microbiology, mineralogy and technology will be employing electron microscopes.

At the Institute for Electron Microscopy at Delft the microscope described here has long been in daily use for investigations in the interest of industry and science.

ELECTRONIC CONDUCTIVITY OF NON-METALLIC MATERIALS

by E. J. W. VERWEY.

537.311.32

Generally speaking, non-metallic materials are electrical insulators. In certain cases, however, conduction — caused by electrons — may take place also in these substances. This article deals with the relation frequently existing in oxides and halogenides between electronic conductivity, photo-electrical phenomena, light absorption and deviations from the stoichiometric composition. From the simultaneous occurrence of all these phenomena some insight is obtained into the manner in which they are brought about. The picture that can be formed is especially worked out for the case of the crystal of potassium chloride, for which many observations and theoretical data are available.

Electrical conductivity in metals is brought about by the movement of free electrons, which are always present. In non-metallic solids, on the other hand, conductivity may be due to two causes: in the first place — as in the case of metals — to the movement of free electrons (electronic conductivity) and in the second place to the movement of ions (electrolytic conductivity). Often both movements take place simultaneously. In this article, however, we will confine our considerations to those non-metallic substances in which the electronic conductivity plays the main part.

The interest in electrotechnics in the phenomenon of electronic conductivity in non-metallic substances is due in the first place to two applications of these substances.

The first relates to the use of all kinds of substances as insulators, thus as media that are required to have the least possible conductivity for an electric current. As a rule the insulating medium is applied for two purposes, either as a dielectric in condensers or as insulating material. In the first case the highest possible capacity is required and, therefore, often a high dielectric constant; as a dielectric titanium oxide (rutile: TiO_2) may for instance be used, the dielectric constant of which in the sintered condition is over 100. In the second case, however, a low dielectric constant is often required and, for instance, organic substances like the well-known polystyrene are used as insulating material. When inorganic substances are indicated because greater chemical and thermal stability are required, materials like porcelain or many of its various modern varieties are used, sometimes in the porous state.

A general answer to the question which material is an insulator and which properties are connected with a low conductivity is therefore of direct technical interest.

In the second application it is rather a high conductivity that is required; in recent years more and more use is being made in electrotechnics of more

or less well conducting materials of a non-metallic nature as resistors, advantage being taken, among other things, of the fact that these substances possess a much higher specific resistance than metals. According to the application envisaged, materials are required with the lowest possible temperature coefficient, sometimes a positive one or in other cases rather a very strongly negative one. The conductivity of these insulators must be purely electronic, because electrolytic conductivity always gives rise to chemical changes at the contacts of the resistor or to polarisation phenomena. In this case it is also often required that the resistance mass must be able to withstand high temperatures, and in such cases inorganic substances are indicated.

The materials referred to above, having a very much lower conductivity than metals but a better conductivity than insulators, are often called semi-conductors. Rather arbitrarily the following distinction is usually made:

Metals: specific resistance $< 1 \, \Omega \, \text{cm}$ (mostly 10^{-5} — 10^{-6} ; sometimes higher).

Semi-conductors: specific resistance between 1 and $10^{10} \, \Omega \, \text{cm}$.

Insulators: specific resistance $> 10^{10} \, \Omega \, \text{cm}$.

All these limiting values for the specific resistance are those applying at room temperature.

Conductivity and non-conductivity appear to be distributed among the elements in such a way that the permanent non-conductors occur in the top right-hand corner of the periodic system. Typical examples of non-conductors are sulphur and diamond. Carbon, however, is on the border line, for its stable modification graphite, is a conductor, at least in certain directions of the crystal. There are also a few intermediate cases occurring in the boundary region, such as selenium, which is a semi-conductor.

These intermediate cases are also found among the compounds. For instance in a series Ag_2O , Ag_2S , Ag_2Se , Ag_2Te we find the conductivity gradually

increasing parallel with the metallic character, which manifests itself, for instance, in increased reflectivity. The problem of electronic conductivity may be identified here with that of the metallic state. Electronic conductivity, however, also frequently occurs in compounds which in other respects are typically non-metallic. In such cases the occurrence of conductivity appears to coincide with a number of other phenomena which throw considerable light upon the process of electronic conductivity and which will be discussed further in this article.

Relation between electronic conductivity, deviation from stoichiometry and light absorption

It appears that in general we have to make a distinction between stoichiometric and non-stoichiometric compounds. Formerly the belief was held that the typical feature of compounds was that they contain the various elements in a simple ratio. In the case of many compounds this is so, but it is not always true for compounds in the solid state. In elementary instruction in chemistry one often uses the familiar example of 56 parts Fe to 32 parts S, forming the solid FeS. It is just for a compound like FeS that it has now been found that it always contains an excess of sulphur, and equally so FeO always contains an excess of oxygen.

These are compounds which cannot exist at all in the stoichiometric proportion, or, to be exact, can only exist in the metastable state. There are, however, also many compounds that may occur both in the exactly stoichiometric proportion and with a different composition. Typical examples are: ZnO, with or without an excess of Zn; TiO_2 , with or without an excess of Ti; Cu_2O and NiO on the other hand sometimes contain an excess of O. In such cases there is a discoloration of the material towards black: ZnO is white, whilst with an excess of Zn it becomes grey; TiO_2 is pale yellow and by chemical reduction becomes dark blue or even black; Cu_2O , in itself red, through oxidation becomes black; NiO, itself pale green, with a small excess of oxygen likewise turns black.

All these substances have the common property that in the stoichiometric form they are good and sometimes even excellent insulators, but that if there is any deviation they get a more or less strong electronic conductivity. TiO_2 , in itself possessing a specific resistance of at least $10^{10} \Omega \text{ cm}$, upon being heated in hydrogen is reduced to $\text{TiO}_{1.75}$ with a specific resistance of $10^{-2} \Omega \text{ cm}$. As a rule electronic conductivity increases with the deviation from the stoichiometric proportions: the specific resistance

of $\text{TiO}_{1.9995}$ is $10 \Omega \text{ cm}$, and that of $\text{TiO}_{1.995}$ is $1.2 \Omega \text{ cm}$.

Compounds are also known in which deviation from the stoichiometric proportion does not immediately result in a blackening of the material but rather in an absorption band in the visible spectrum. BaO (white) can incorporate about one per cent Ba in excess and thereby turns red. KCl heated in K vapour incorporates a small excess of K (at most a few hundredths percent) and becomes violet. Similarly NaCl may be given an excess of Na, the crystals thereby turning yellow. In these cases the electronic conductivity resulting from the deviation from the stoichiometric composition is not noticeable at room temperature; slightly higher temperatures are required to be able to observe a decided difference compared with the conductivity of the pure compound. For the purpose of our considerations, however, this makes no essential difference.

We therefore observe the simultaneous occurrence of three phenomena: deviation from stoichiometry, absorption of visible light and conductivity caused by electrons.

In how far these phenomena are related to each other can be illustrated by the following experiment (fig. 1), which was first described by Stasiw, one

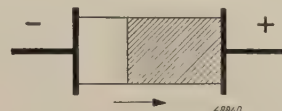


Fig. 1. When a violet-coloured KCl crystal is placed between two electrodes at a temperature of about 500°C the colouring moves out of the crystal. The arrow indicates the direction in which the boundary between the violet (shaded) part and the colourless part of the crystal moves.

of Pohl's co-workers, in whose laboratory an intensive study was made of the phenomena occurring with this kind of discoloured crystals.

Two electrodes are applied to the two opposite faces of a KCl crystal which has been made violet by the incorporation of additional potassium atoms and which is kept at a temperature of 500°C (KCl was chosen by preference for this experiment because the violet colour can be clearly observed). In the first place it is found that the conductivity of the coloured crystal is much greater than that of the non-coloured crystal at the same temperature. If the contact between the crystal and the negative electrode is rather faulty, then, on the current being passed through, a non-coloured zone is seen to form near the negative electrode, which zone gradually extends over the whole crystal. The discoloration of the crystal, which at this temperature has become

deep blue, is seen to move through the crystal and after some time disappears into the electrode; this is accompanied by a gradual reduction in conductivity. After completion of the experiment the crystal again becomes stoichiometric, colourless KCl, and the extra conductivity has disappeared.

This remarkable experiment, the interpretation of which will be reverted to later, demonstrates very clearly that the conductivity created by the excess of potassium and the violet colour are closely related to each other.

Phenomena connected with irradiation

There is a second group of phenomena which throws some light upon the mechanism of the conductivity of these substances, *viz.* the photo-conductivity and the photo-chemical changes taking place under the influence of rays of different wavelengths.

The phenomenon of photo-conductivity occurs in all sorts of substances; it implies that when a substance is irradiated with light of certain wavelengths an originally non-conductive substance becomes conductive, or rather that the conductivity of the substance is increased by the irradiation. It occurs for instance with the KCl crystals just mentioned (which by heating in K vapour have turned violet) when the crystals are irradiated with light that is absorbed by the violet coloured crystal. The violet colour of this KCl is caused by an absorption band in the yellow part of the spectrum with a maximum at about 5600 Å. The phenomenon of photo-conductivity therefore occurs under irradiation with yellow light of about this wavelength.

From this behaviour it is to be concluded that under the influence of light of this wavelength electrons which before the irradiation were bound to certain places inside the crystal are released. The energy with which these electrons are bound to those places and which, therefore, has to be supplied to release an electron follows directly from the wavelength of the active light. This wavelength λ (frequency ν) corresponds to the energy eV supplied to the electron by the absorption according to the formula

$$eV = h\nu, \text{ or: } V(\text{volt}) = \frac{hc}{e\lambda} = \frac{12390}{\lambda(\text{\AA})},$$

in which V is the voltage through which a free electron has to pass in order to accumulate the energy required (h , c and e indicate respectively Planck's constant, the velocity of light and the charge of the electron). From this we calculate for the band at 5600 Å in coloured KCl that the elec-

trons are bound with an energy of 2.2 eV (the electron energy is expressed in "electron volts" eV by indicating the voltage V that corresponds to the energy eV).

With the excess of potassium we have thus introduced additional electrons in the crystal, which may be in two states; the state in which they are still bound and a state with 2.2 eV higher energy in which they have freedom of movement and may cause conduction. Further we have discovered two different means of raising the electrons from the lowest state of energy to a state of higher energy, namely by supplying thermal energy (raising the temperature to about 500° C suffices to make the conductivity perceptible) and by supplying radiation energy¹⁾.

The occurrence of conductivity may also be observed in crystals of potassium chloride coloured violet by photo-chemical means. The same violet colour that can be obtained by heating in potassium vapour is also obtained when KCl is irradiated, for instance, with X-rays or with ultra-violet light. KCl crystals coloured in this way, however, behave in a slightly different way than do the crystals with an excess of potassium, but, nevertheless, the electrons bound with an energy of 2.2 eV can still be released, *i.e.* on heating or by irradiation with yellow light conductivity is observed. The photo-chemically coloured crystals, however, differ from the others in one respect: upon irradiation or heating the violet colour very soon fades away and the crystal returns completely to its original state. Apparently under this treatment the electrons find an opportunity to return to their original positions. Photo-chemically coloured crystals are therefore in a metastable state; when by the addition of energy a return to the stable state is made possible, this takes place by transport of electrons inside the crystal lattice.

The discoloration with X-rays is not really a purely photo-chemical process but is caused mainly by the secondary electrons released in the substance. As a matter of fact the same result can be obtained by the irradiation of a KCl crystal with electrons. A recent example of this phenomenon is found in the cathode ray tubes where the screen consists

¹⁾ For the sake of clarity, on a few minor points we take the liberty of representing the position in a somewhat too simple manner. One of these points is that the energy to be supplied by the thermal movement in order to release an electron from the bound state is quantitatively not equal to the radiation energy required for the same purpose, but in general about half of that. Another simplification is that in reality the above-mentioned transition to the state with a higher energy probably takes place *via* an intermediate state; the energy of this intermediate state, however, is only very little less than that of the final state.

of a thin layer of vaporised alkali-halogenide; in this kind of tube advantage is taken of the fact that this layer has the property of discoloring more or less permanently at the places where it is struck by the electron beam.

Further considerations about conductivity and light absorption

At room temperature the mean energy of the thermal movement of an ion or a free electron amounts to 0.04 eV. In coloured KCl the transition from a weakly bound electron into a free electron involves many times this amount, so that the chance of this transition being brought about by thermal movement is practically nil at room temperature. When the temperature is raised, however, this chance increases very rapidly, with the result that at a temperature of about 500° C a noticeable number of electrons are released and conduction can be observed in the manner just described.

The fact that non-coloured KCl is an insulator and at about 500 °C still does not yet show any noticeable conductivity for electrons may be related in the same way to the absorption in the ultra-violet part of the spectrum. As already remarked, this absorption brings about a discoloration of KCl, *i.e.* the release of electrons in non-coloured KCl, so that the electrons released in this way may at least partly be caught in the same positions which are also taken up by the additional electrons of KCl heated in K vapour. The energy required to release an electron inside the crystal of non-coloured KCl is thus much greater than that required in the case of coloured KCl; the corresponding maximum in the absorption band lies at 1310 Å, from which we can calculate, with the aid of the conversion formula previously used, a corresponding energy of 9.4 eV. The probability of this large amount of energy being supplied by the thermal movement is again very much less than in the case of coloured KCl, and consequently even at a higher temperature the electronic conductivity in non-coloured KCl is practically negligible.

In the case of the oxides previously mentioned, where there is a deviation from the stoichiometric composition the colour is mostly black, absorption thus occurring throughout the whole of the visible spectrum; moreover, this absorption extends as a rule far into the infra-red. This means that these substances contain electrons which are bound to certain places with very low energies. This is in accordance with the fact that already at room temperature a noticeable number of electrons are

released and electronic conductivity takes place in these substances.

Interpretation in connection with the crystal structure

Potassium chloride forms an ion lattice with positive metal ions and negative halogen ions. The occurrence of ions is related to the tendency of all atoms to surround themselves with completed shells of electrons. Potassium, which has one electron in the N shell, loses this, whilst chlorine, with 7 electrons in the M shell, supplements it to eight electrons.

To take one electron from a potassium atom an energy of only 4.1 eV is required; to remove a second electron would cost 44 eV, because it would then have to be taken out of the completely filled M shell. The ionisation energy of the chlorine atom is 13.1 eV, which is much greater than that of the potassium atom. Apparently there is little tendency on the part of the chlorine atom to give off an electron out of its M shell. Rather it would be inclined to supplement this M shell, which is already almost full; in taking up the one missing electron an energy of 3.6 eV — called the electron affinity — is gained, whereas in the case of potassium no gain in energy can be attained by the taking up of an electron.

Although the ionisation energy of potassium is still 0.5 eV greater than the electron affinity of chlorine, yet the formation of an ion lattice is advantageous from the point of view of energy. As the positively and negatively charged ions approach each other a large amount of potential energy is released which makes the ion lattice more favourable from the point of view of energy in comparison with an atom lattice. This energy is usually written in the form

$$E = M \cdot \frac{e^2}{r},$$

in which r is the distance between adjacent ions and M is a numerical factor depending upon the type of lattice (the so-called Madelung constant). For the lattice type of KCl $M = 1.75$. Further $e = 4.8 \cdot 10^{-10}$ e.s.u. and $r = 3.14 \cdot 10^{-8}$ cm, so that

$$E = 12.8 \cdot 10^{-12} \text{ erg} = 8.0 \text{ eV.}$$

As a consequence of the formation of the lattice 8.0 eV is thus gained per pair of ions, and since the formation of the ions out of the atoms costs only 0.5 eV the ionised state in the crystal is still more favourable by 7.5 eV than that of the free neutral atoms.

It is therefore readily seen that an extra electron added to this lattice built up of ions will be able to move through the lattice quickly and that, therefore, the same will be the case with an electron originating from the lattice itself, once it has been released from the ions making up the lattice. Let us suppose that at a given moment this extra electron is at the position of a K^+ ion and forms with it a potassium atom. Owing to the large radius of the potassium atom (or in other words owing to the expansion of the trajectory of this electron), the electron will not merely stay in the cavity of the K^+ ion but will penetrate into the neighbouring chlorine ions. Fig. 2, in which the dimensions

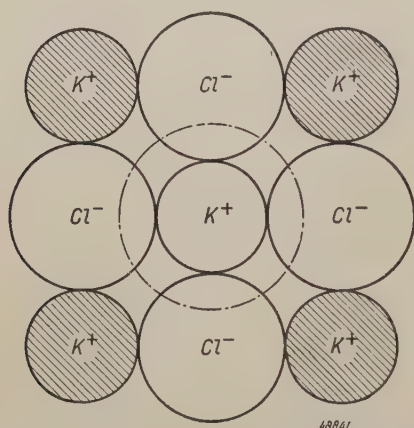


Fig. 2. Structure of a KCl crystal with potassium ions and chlorine ions in their true position and size. The dot-dash circle indicates the space which a potassium atom would occupy in the position of a K^+ ion.

of the ions and of the potassium atom are represented in their true proportions, shows that in the course of its normal movement the electron will frequently reach a point half-way between two adjacent K^+ ions, so that it will easily pass on from one ion to the other. A small electrical field strength is sufficient to cause a directed movement of the electrons, which manifests itself in the form of electrical conductance.

It is, however, likewise readily to be understood that in a stoichiometric KCl lattice it will be very difficult to release an electron from one of the ions of which the lattice is composed. For this only a Cl^- ion can be considered, since the ionisation of the K^+ ion costs 44 eV. It still requires rather a lot of energy to remove an electron from a Cl^- ion, owing to the fact that, as we have seen, the electrostatic binding forces stabilise the state where the binding is brought about by ions. Consequently we have to supply not only the energy required to release an electron from Cl^- (3.6 eV; see above) but also that required to overcome the electro-

static forces originating in the surrounding lattice, whereby the attraction from the adjacent positive K^+ ions predominates. The potential energy of an electron in the position of a Cl^- ion as a result of the electrostatic lattice forces is in fact $-M \cdot e^2/r$ or -8.0 eV. In stoichiometric KCl, where the electrons derived from the potassium atoms are bound by just as many chlorine atoms, and in general in ion lattices of stoichiometric composition, the energy required to release electrons from the negative ions is therefore considerable, with the result that these substances are not coloured and show no electronic conductivity. We need not, it is true, supply the full $3.6 + 8.0 = 11.6$ eV²⁾ to remove the electron from the Cl^- ion out of the field of the surrounding ions and to let it move freely through the lattice.

Only 9.4 eV suffices, corresponding to the aforementioned absorption at 1310 Å. Nevertheless, this is still a considerable amount of energy.

The consideration given here regarding the various states of energy in which the electrons may exist in a KCl crystal might also be cast in a somewhat different form, seeking a closer connection with the quantum-mechanical theory of the solid substance. According to quantum-mechanics the energy of the electron in a solid cannot *a priori* have all possible values. Certain energy zones are "forbidden" for electrons while others are "allowed", the latter being called "energy bands". The situation is diagrammatically represented in fig. 3a, where the different shading of the three energy bands indicates that the two lower ones are "occupied" by electrons, whereas the top one is "unoccupied". This way of expressing the picture calls for some explanation. One must imagine an energy band as being formed by a number of energy levels lying very close together. According to quantum-mechanics any two electrons forming part of the same system (e.g. a piece of solid substance) can never have exactly equal energy. In our case, expressed somewhat differently, this means that each level in an energy band can only be occupied by at most one electron. When all levels of an energy band are occupied by electrons the band is said to be "occupied" or "full". If no level in a band is occupied by electrons one speaks of an "unoccupied" or "empty" band. Of course an energy band may also be partly occupied (see fig. 3b). The conductivity of a substance depends upon the manner in which the energy bands are occupied by electrons. If all the energy bands are either fully occupied or fully unoccupied then as a rule the substance is an insulator (fig. 3a); if on the other hand there is an energy band only partly occupied then we have a conductor (fig. 3b). This is closely related to the fact that in the former case a

²⁾ For instance owing to the fact that the surroundings of the remaining "positive hole" adjust themselves to the change, by which process a certain polarisation energy is gained. Further, in the estimation it is presumed that in the conductive state the electrons are bound with a binding energy equal to zero because on their way through the lattice they are alternately located near K^+ and Cl^- ions. In reality these electrons still have a certain interaction with the lattice. The whole estimation given above is only approximative.

finite, in some cases large energy increase is necessary to pass an electron to a higher unoccupied level, whereas in the latter case a practically infinitely small energy increase is needed. The position and width of the energy bands depend upon the chemical composition and crystal structure of the substances³⁾.

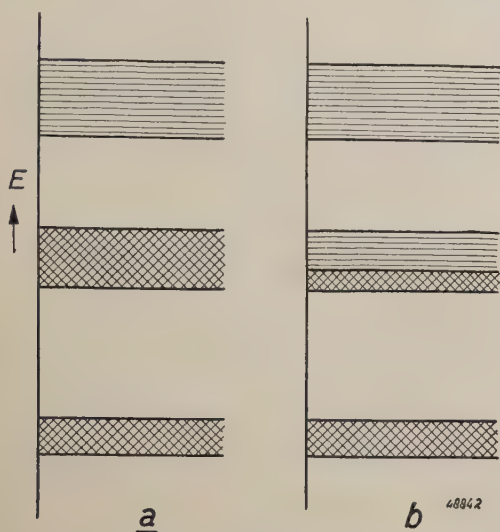


Fig. 3. Diagrammatic representation of the energy bands of electrons in a solid. The values of the electron energy E lying between the values corresponding to the bands are "prohibited". All energy bands situated below the three drawn here are fully occupied by electrons and all energy bands above those drawn are unoccupied. In case *a*) there are no partly occupied energy bands; the two lower bands are entirely occupied (cross-hatched), the top band is unoccupied (shaded); the material is an insulator. In case *b*) there is one partly occupied energy band (the middle one): the material is a conductor (e.g. a metal).

In this article our considerations have been based on the simplified picture that the KCl lattice is an ion lattice. In this picture the electrons that can be released at the lowest possible cost of energy belong to the Cl^- ions. Also according to quantum-mechanics such a picture is very near the truth for a strongly polarised substance such as KCl. The energy state of the electrons bound in the Cl^- ions therefore corresponds to that of the highest occupied band of KCl. The state of electrons which by the addition of an amount of energy of at least 9.4 eV have been released from these Cl^- ions and are more or less free to pass through the lattice corresponds to the state of the lowest unoccupied band. The position of the bands in stoichiometric KCl is diagrammatically represented in fig. 4*a*. At absolute zero temperature all electrons are in their fundamental state, and since all Cl^- ions are complete the band *A* is fully occupied. Consequently there are no electrons in the excited state and the energy band *B* is completely "empty". The distance between these two bands is so great that even at room temperature only very few electrons pass from band *A* to band *B*.

In potassium chloride containing an excess of potassium the extra electrons derived from the excess potassium atoms have to find room in other

positions where they are naturally not so strongly bound as in the Cl^- ions. Regarding the nature of this state of binding, which must exist both in KCl discoloured through heating in potassium vapour as well as in photo-chemically discoloured KCl, the following picture has recently been formed (J. H. de Boer, Mott).

If an excess of potassium is incorporated in a KCl crystal this is probably not brought about by the K^+ ions formed in the lattice finding their place between the normal lattice ions. Schottky and others have made it plausible that also a stoichiometric KCl crystal contains a number of lattice defects in the form of unoccupied places; K^+ ions and Cl^- ions are then missing in equal numbers (fig. 5*a*). In non-stoichiometric KCl with an excess of potassium there are more Cl^- ions missing than K^+ ions, the excess of K^+ ions being located in normal lattice position; the extra electrons from the atoms having been introduced into the crystal will preferably occupy the position of the missing Cl^- ions (fig. 5*b*). The above mentioned energy of 2.2 eV is the energy binding the extra electrons of these "lattice defects".

When a stoichiometric KCl crystal is irradiated with X-rays or ultraviolet light the result is that an electron is taken away from some of the Cl^- ions; these released electrons will again preferably occupy the places of the missing Cl^- ions, so that

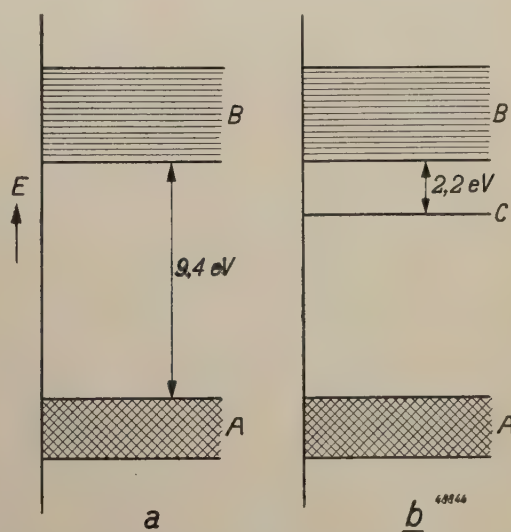


Fig. 4. Simplified diagram of energy bands:

a) in the case of stoichiometric KCl with a crystal lattice without defects and therefore not coloured (band *A* is entirely occupied, band *B* is empty);

b) in the case of KCl with lattice defects, i.e. holes (cf. fig. 5). In this case there is a new and very narrow energy band *C*, drawn here as a sharp level; if the KCl is non-stoichiometric or has been irradiated with X-rays then band *C* is partly occupied and the crystal has a violet colour. Through exposure to yellow light or through heating, electrons may then pass over from *C* to *B*, the latter then becoming partly occupied, thus making coloured KCl a conductor.

³⁾ Considerations about energy bands have also been applied in this periodical for explaining the behaviour of blocking-layer photocells (see Philips Techn. R. 8, 56, 1946) and of luminescent materials (see Philips Techn. R. 6, 349, 1941).

the binding energy of these electrons is again 2.2 eV (fig. 5c).

The difference between KCl coloured by X-rays and KCl coloured by heating in potassium vapour consists in the fact that the former contains in addition to ions and extra electrons also neutral chlorine atoms, namely in the places where a chlorine ion of the lattice has given off an electron under the influence of the irradiation. For this reason only the colour of KCl that has been coloured by X-rays

One may look for some connection with the quantum-mechanical considerations introduced above concerning the energy bands in solids also of non-stoichiometric KCl. A diagram of the energy bands of KCl crystal coloured through an excess of potassium is given in fig. 4b. At absolute zero temperature all the extra electrons in such a crystal occur in the energy band *C*. This energy band may also be present in stoichiometric KCl but there is quite unoccupied (unless electrons pass into it from band *A* through irradiation with X-rays). The only condition for the presence of band *C* is the presence of unoccupied positions in the crystal lattice of KCl.

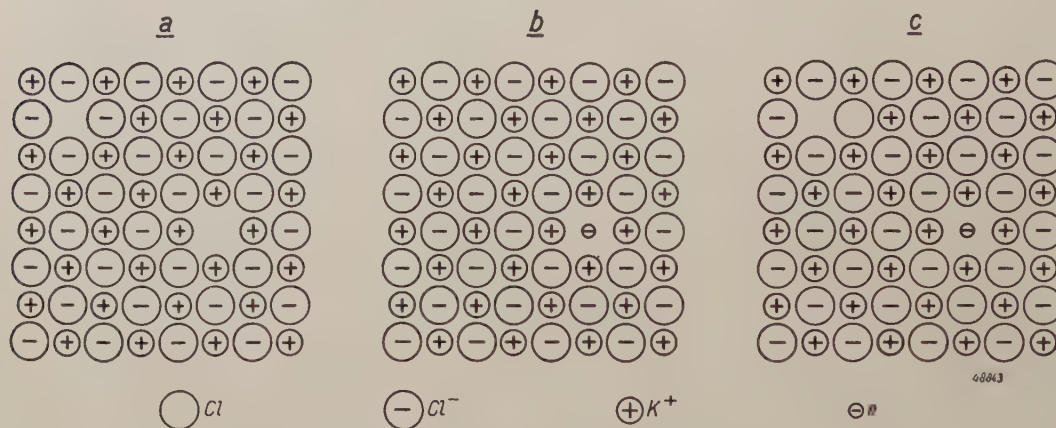


Fig. 5. Diagrammatic representation of the structure of KCl:

- stoichiometric KCl with lattice faults, i.e. unoccupied positions (in the figure one K^+ hole and one Cl^- hole have been drawn).
- KCl with an excess of potassium; the atom introduced into the lattice separates into a K^+ ion, so that the K^+ hole becomes occupied, and an electron (e) which is situated in the position of the Cl^- hole.
- Stoichiometric KCl irradiated with X-rays or ultra-violet rays; the electron thereby separated from a Cl^- ion (at the K^+ hole) comes to lie in the Cl^- hole. In the process of discoloration (due to excess of potassium or to irradiation) as a rule the holes shift to different positions. For the sake of clarity this has not been taken into account in the figure.

will fade again when heated or irradiated with light from its absorption band of 5600 Å. It is only in this case that the electrons that have again been released have an opportunity to move into the free places and again form chlorine ions.

From the estimations made in the foregoing it follows that the positions where Cl^- ions are missing are indeed well suited to hold an electron. As a matter of fact for the potential energy of an electron in the centre of the Cl^- cavity we again find $-1.75 \cdot e^2/r = -8.0$ eV. Actually the potential energy in the cavity is very much less negative, owing to the polarisation of the immediate surroundings (cf. note ²). Moreover, the electron retains some degree of movement, so that it does not always remain in the centre of the cavity, and possesses a certain kinetic energy. Consequently the resulting energy lies approximately 2.2 eV lower than that of the conducting state.

Since the extra electrons in coloured KCl are rather great distances apart and thus influence each other very little, they will all have approximately the same energy, in other words band *C* is very narrow, almost a sharp level. The photoconductivity caused by irradiation with yellow light or the conductivity arising from the heating of the coloured KCl to about 500° C is interpreted in this system of bands as a consequence of the fact that a number of electrons are passed from band *C* into the unoccupied "conductivity band" *B*.

The estimations given here for the energy of the various states of the electrons in the lattice of the alkali halogenides could be replaced by more accurate calculations, so that it may be said that our knowledge and understanding of the conductivity phenomena in the alkali halogenides forms a more or less complete picture. Regarding other cases, such as the technically so much more important oxide semi-conductors, the theory has not been so far developed, because in these cases

quantitative calculations are much more difficult. Nevertheless, the phenomena are very similar.

We can now look somewhat more closely into the experiments described in the foregoing, where the colouring of non-stoichiometric KCl disappeared from the crystal under high temperature.

At the temperature of the experiment, owing to thermal movement electrons are continuously changed from the weakly bound state (in "Cl⁻ holes") into the conductive state. Under the influence of the electric field they will in course of time assume on the average a directed movement towards the positive electrode. Sometimes they will thereby again be caught in empty "Cl⁻ holes", and after a certain time again be released, and so on. Upon reaching the positive electrode they will ultimately disappear into it. If the contact with the negative electrode is a very good one then electrons will again be supplied there through transition from the electrode to the crystal. The passage of the current can then only be registered by the current intensity observed. The crystal remains coloured.

The phenomena described in the experiment occur when the contact at the negative electrode is incomplete, thereby making it difficult for electrons to pass from the electrode to the crystal. A thin layer of colourless KCl is then formed between the electrode and the coloured crystal. This layer has a much smaller conductivity than a coloured KCl, though it is not entirely zero owing to the fact that at the temperature of the experiment KCl has a small electrolytical conductivity due to the ions or, more probably, the "ion holes" already possessing a certain freedom of movement. The current, through gradually decreasing with increasing thickness of the non-coloured part of the crystal, is thus maintained by a stream of ions, the K⁺ ions collecting on the surface on the crystal at the negative electrode, where they are discharged and form free potassium. Since the whole crystal remains electrically neutral everywhere, the quantity of separated potassium is equivalent to the quantity of electrons passing into the other electrode. The whole phenomenon therefore resolves itself into the disappearance of the excess of potassium out of the crystal owing to electrons being given off to one electrode and the equivalent quantity of K⁺ ions to the other electrode. It is thereby observed that the extra electrons and the violet colour are also locally connected one with the other. It may therefore be said that this experiment makes the movement of the conductivity electrons visible.

Practical applications

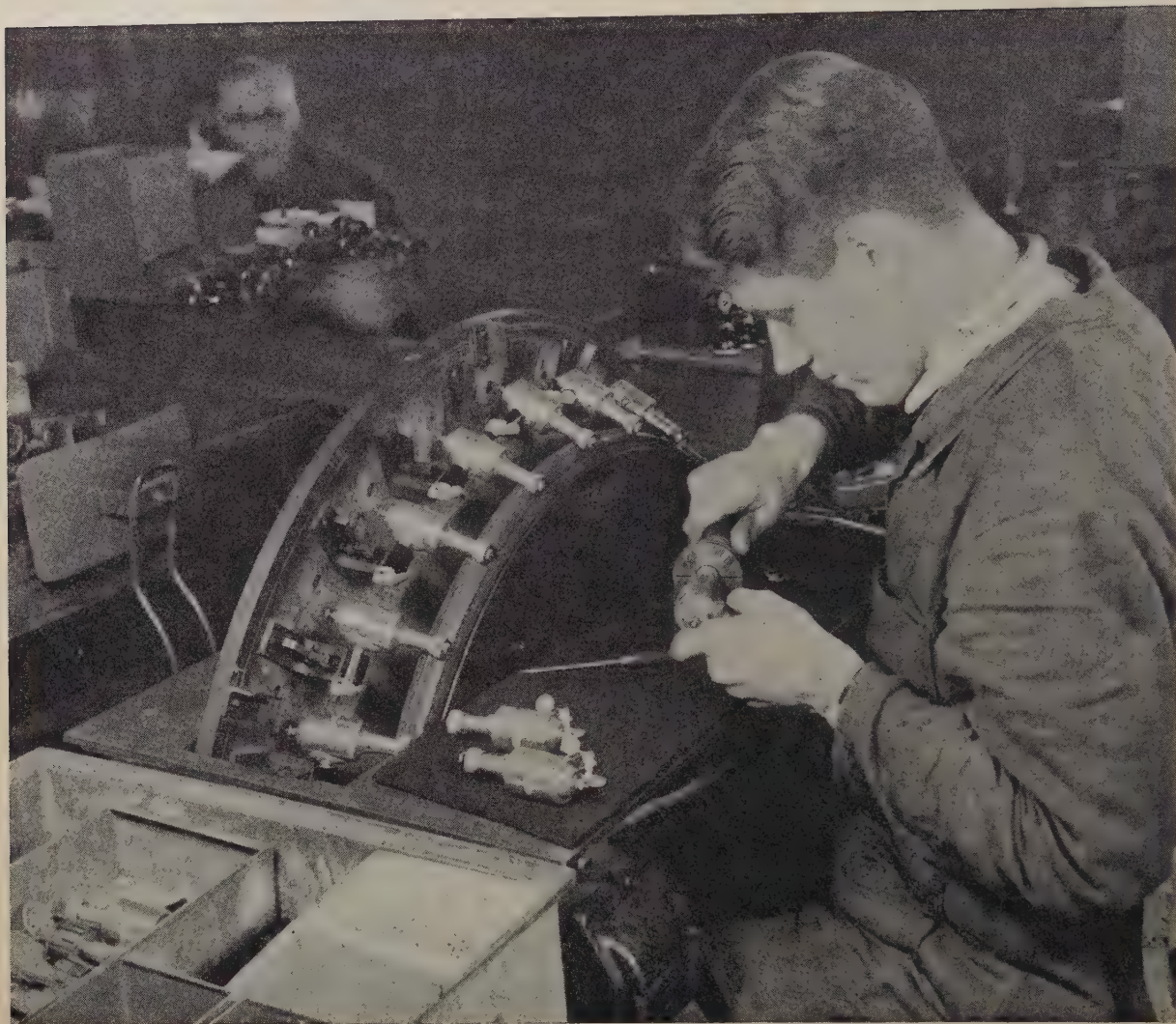
As regards insulating materials the choice made in practice is such that one does not as a rule use materials in which any deviation from the stoichiometric composition readily occurs. If for some other reason one should, nevertheless, decide to

employ such a material one must be careful to take the possibility of such deviations into account. As an example we might mention titanite oxide used as dielectric for condensers, which material has to be sintered by heating to a high temperature. One of the precautions that has to be taken in heating it is that the gas atmosphere and the heat treatment must be such that the final product shows no trace of a reduction to oxides of titanium of a lower valency. The losses in a high-frequency alternating field (dielectrical losses) in particular are very sensitive to the slightest deviation from the stoichiometric composition of TiO₂. Quantities of titanium of a lower stage of oxidation which cannot be detected by chemical means may increase the normal value of the dielectrical losses (tg δ several times 10⁻⁴) by a factor of 10 or 100. Owing to the relatively easy reduction of TiO₂ one is somewhat limited also in the use of this material and care must be taken to avoid any contact with organic materials at temperatures higher than a couple of hundred degrees centigrade.

Just as the introduction of potassium atoms (with one electron more than the K⁺ ions) in KCl involves the introduction of conductivity electrons, with an excess of titanium we introduce into the crystal lattice of TiO₂ in some way or other weakly bound electrons which cause the conductivity phenomena. One may imagine that part of the Ti⁴⁺ ions is changed into Ti³⁺ ions by taking up an electron; perhaps, however, this conception is too simple and the state of these bound electrons may be compared to that in coloured KCl.

For the resistance materials mentioned in the beginning of this article use is in fact made of these very properties of some oxides. These materials often contain, in addition to oxides having insulating properties, also a certain percentage of oxides which are easily brought to the conductive state or usually are already in that state. Apart from titanite oxides also iron oxides for instance are used, the conductivity of which is due to the presence of Fe²⁺ ions, which again contain one electron more than the Fe³⁺ ions likewise present. We hope to deal with these materials more fully in another article to be published shortly.

RUNNING IN CYCLE DYNAMOS



The final processes in the manufacture of cycle dynamos are the running in, the adjustment of the driving wheel (reducing the axial play to the minimum) and the mechanical and electrical testing.

A number of dynamos, e.g. 24, are run in simultaneously. These are secured to a large disc that can be turned into various positions. They are all driven by a single wheel with a rubber tire of the same diameter as an ordinary bicycle wheel and rotated at a peripheral speed corresponding to the speed of a bicycle of about 15 km per hour. When the dynamo is taken off the disc to be adjusted, another is put in its place, while at the same time the disc is turned to the next position. In this way every dynamo is run in for about 20 minutes, the adjustment taking only about 1 minute.

INSTALLATIONS FOR IMPROVED BROADCAST RECEPTION

by P. CORNELIUS and J. van SLOOTEN.

621.396.666:621.396.677.1

Broadcast reception is often made unsatisfactory through various causes, the most important of which are: selective fading effect and interference from transmitters on neighbouring wavelengths. Thanks to the fact that fading effect seldom occurs at different places simultaneously, its unpleasant consequences can be successfully counteracted by setting up receivers some distance apart (diversity reception) and connecting to the loudspeaker(s) only that one where the reception happens to be best at the moment. An apparatus has been worked out which brings this about automatically. In practice it appears that three receiving stations about 1 km apart are sufficient. Interferences from other transmitters can be counteracted by applying directional reception with the aid of a frame aerial, preferably in combination with a normal antenna. Thanks to the freedom from disturbances thereby attainable, the bandwidth of the receiving set can readily be increased, thus improving the quality of the sound. Diversity reception and directional reception can easily be combined. The former, however, can only be considered for installations serving a large number of listeners.

Introduction

For the reception of the transmission from a broadcasting station to be such that the reproduction of the music and the spoken word satisfies reasonable demands, it is necessary that the field of the transmitter has a certain minimum intensity. This minimum field strength is determined theoretically by the noise voltages¹⁾ present in every receiver and by the atmospheric and local disturbances that almost invariably occur.

Anyone who regularly listens in, or tries to do so, to the programmes of stations a great distance away may find that in practice there are several factors which often make a reasonably satisfactory reception impossible, even though the field strength is greater than the minimum just mentioned. The causes of this are:

- 1) the so-called selective fading and
- 2) interferences from other transmitters.

In the following pages it will be explained how the effect from both these causes can be eliminated by employing special methods of reception. Unfortunately these methods are still too complicated and too expensive to be applied on a large scale by individual listeners. On the other hand they are sufficiently simple and reliable to be used wherever one is prepared to go to some expense to get good reception of remote stations, as may be the case in large establishments, such as hospitals for instance, where it is desired to distribute the broadcasting programmes to a large number of listeners.

We will now discuss separately the two above-mentioned causes of unsatisfactory reception and

indicate how they can be rendered harmless. Although these methods differ for the two causes, it will be seen that there is in principle no objection against their being combined into one system of reception.

Selective fading and its counteraction

It is a well known phenomenon that after night-fall several distant broadcasting stations can be received which during daytime can hardly be heard at all. This is connected with the fact that the propagation of radio waves, particularly over long distances, depends a great deal upon the state of the ionosphere. The ionosphere is a region at a great altitude in the atmosphere where, owing to the influence of the ultra-violet radiation from the sun, ionisation takes place and one or more electrically conducting layers are formed²⁾. When radio waves of a sufficiently large wavelength strike this layer at a favourable angle of incidence they are diffracted back to earth. In the course of a period of 24 hours the condition of the ionosphere shows a periodical change owing to variation in the radiation from the sun. As a consequence in the evening conditions are favourable for bridging long distances, particularly by means of the wavelengths of 200-600 metres.

Besides these diurnal variations (and others of longer periods) there are much more rapid and less regular changes in the state of ionisation. These changes affect the phenomena of interference arising from the fact that the waves from a transmitter may reach a receiving aerial in two ways,

¹⁾ Regarding noise in receivers see e.g. Philips Techn. R. 3, 189, 1938.

²⁾ A comprehensive article on the radio investigation of the ionosphere appeared in Philips Techn. R. 8, 111, 1946.

either one way *via* the earth's surface and one *via* the ionosphere, or both *via* the ionosphere. Owing to the irregular changes just mentioned in the ionic density of the ionosphere the maxima and minima of interference are continuously changing, and this manifests itself in a receiver as a continuous variation of the field strength. This is what is known as "fading"; the phenomenon occurs on wavelengths of 200-600 metres only at night and during twilight, because it is only then that the ionosphere affects reception.

As long as the paths along which the waves reach the aerial show differences in length only of the order of the carrier wavelength, the fading effect occurs simultaneously for the carrier wave and for the side bands. As a consequence of the automatic volume control usually present in a radio receiver this is little noticed. It becomes a more serious matter when the said differences in the paths assume magnitudes of the order of wavelength corresponding to the audio frequencies of the modulation. Then the fading is highly selective, that is to say it depends to a high degree upon the modulation frequency. Consequently for the components of the frequency spectrum emitted by a broadcasting station the transition at a certain moment may be greatly different. This is apt to lead to serious distortion of the signal, and particularly so when the field strength of the carrier wave is too low compared with that of the side bands, resulting in so-called over-modulation. This is the cause of speech so often becoming quite unintelligible.

This undesirable phenomenon of selective fading can in principle be reduced in two ways. One method consists in artificially amplifying the carrier wave in the receiver, but, apart from the fact that this is technically rather complicated and expensive, it has proved to be inadequate, though undoubtedly the average quality of reception can thereby be appreciably improved.

The second method, which is simpler and also more satisfactory, is based on the fact that the fading varies considerably not only with time but also locally; it is found that receivers set up some distance apart and tuned in to the same transmitting station are seldom troubled with selective fading at the same time. The principle of this method, therefore, lies in a mutual comparison of the signals in two or more such receivers and an automatic selection of the best signal. This is known as diversity reception.

Experiments carried out in Philips Laboratory in 1939 demonstrated, *inter alia*, that by a combin-

ation of two receivers about 800 metres apart the reception of a transmitting station greatly subject to selective fading (wavelength 450 m) was so much improved that little trouble remained. When the number of receiving stations was extended to three, at a greater distance apart (about 2 km), selective fading was scarcely noticeable at all. The quality of reception was thereby improved from "very badly disturbed" to "very good".

Although a combination of diversity reception with the aforementioned carrier wave amplification is possible and theoretically may yield still further improvement, in our experiments there was little evidence that such is desirable, a satisfactory improvement being already possible without any such technically objectionable complication.

After these general remarks we will now describe briefly how the method of diversity reception was carried out technically. In each of the three receiving stations (I, II and III in *fig. 1*) there were two

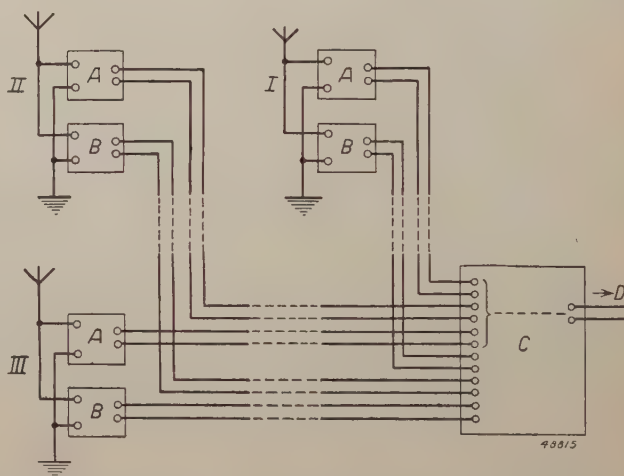


Fig. 1. Block diagram of a receiving system practically free from selective fading. I, II and III are separate receiving stations (at distances of the order of 1 km apart), each with two receiving sets, A and B, the outputs of which are connected *via* transmission lines with a switching apparatus C at a central point. The receivers A are high quality sets. At D a loudspeaker or l.f. amplifier is connected. Of the three sets A, the one where the greatest field strength of the desired transmitting station is received is connected to D by means of C. This field strength is measured by the simple receivers B, which supply the result to C in the form of a direct voltage.

receivers, a special broadcasting receiver (A) of high quality and yielding a l.f. signal of about 1 volt, and a simple broadcasting receiver (B) from which a rectified voltage was drawn to serve as measure for the local signal strength. (Of course both voltages could also be taken from a single receiver specially designed for the purpose.) The two voltages — the l.f. signal and the rectified voltage — were carried through conductors of the local telephone network to a central receiving point,

where a switching unit (C) provided with relays was set up, connecting to the output terminals whichever l.f. signal corresponded with the highest rectified voltage, thus the one coming from the receiving station where the field strength happened to be greatest at the moment. A schematic diagram of this switching unit is given in *fig. 2*.

Briefly this switching system works as follows:

Three triodes (or pentodes), T_1 , T_2 and T_3 , each contain in the anode circuit the coil of a relay (Re_1 , Re_2 , Re_3) and an anode resistance (R_{a1} , R_{a2} , R_{a3}); the triodes are fed *via* a common cathode resistance R_k from a voltage source the poles of

of the valves have a high anode current and one a low anode current. There can only be stability in a condition where one anode current has the full value and the other two are as good as nil³⁾. Which valve receives anode current is determined by the voltages which the receivers B (*fig. 1*) pass to the terminals B_I , B_{II} , B_{III} and which increase with the local field strength of the transmitting station required. Suppose that at a certain moment the field strength is greatest at station II . The voltage on terminals $+B_{II}$ is thus higher than that on $+B_I$ and $+B_{III}$, and *via* the diode D_2 it passes almost entirely also across the resistance R_d , thereby preventing any current across the diodes D_1 and D_3 from the lower voltages B_I and B_{III} . Therefore $+B_{II}$ has approximately the potential 0 with respect to $-b$; $+B_I$ and $+B_{III}$, however, have negative

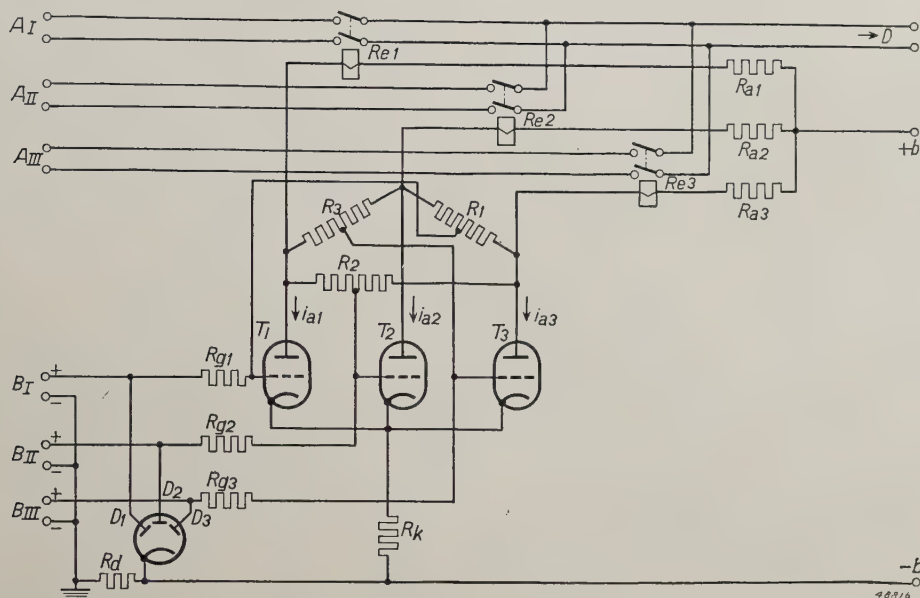


Fig. 2. Diagram of the switching apparatus *C* of *fig. 1*. The lines from the receivers *A* (*fig. 1*) are connected to the pairs of terminals A_I , A_{II} , A_{III} , one of which can be connected *via* the contacts of relays Re_1 , Re_2 or Re_3 to the output terminals *D*. Which pair of terminals is connected to *D* depends upon the direct voltages on the terminals B_I , B_{II} , B_{III} , which are connected to the receiving sets *B* (*fig. 1*). R_k = cathode resistor for automatic negative grid bias of the triodes T_1 , T_2 , T_3 ; R_{g1} , R_{g2} , R_{g3} = current adjusting resistors; $+b$, $-b$ poles of the anode voltage source.

which are $+b$ and $-b$. Between the anodes there are three equal, centre-tapped, delta connected resistances (R_1 , R_2 , R_3); the grids are connected to the middle of the "opposite sides". Assuming for the moment that the pairs of terminals B_I , B_{II} , B_{III} are short-circuited, then the grids are at the same time connected to $-b$ *via* the equal resistances R_{g1} , R_{g2} , R_{g3} and the resistance R_d . Considering the symmetry of the circuit, one would expect the three anode currents (i_{a1} , i_{a2} , i_{a3}) to be equal. This situation, however, is unstable, as may readily be understood when supposing, for instance, that i_{a1} exceeds somewhat i_{a2} and i_{a3} ; in R_{a1} and the coil of Re_1 there will then be a greater voltage loss than in R_{a2} , etc., the anode of T_1 gets a lower potential than the anodes of T_2 and T_3 , the centres of the resistances R_2 and R_3 (to which the grids of T_2 and T_3 are connected) likewise drop in potential, so that i_{a2} and i_{a3} will decrease, thus increasing the initial asymmetry, and so on. In a similar manner it may be seen that it is not possible either for a situation to continue to exist where two

potentials. As a consequence T_2 will carry anode current, while T_1 and T_3 are dead. Thus Re_2 connects the receiver *A* of post *II* with *D*, while the receivers *A* of posts *I* and *III* remain out of circuit. If later on post *III* should have the greatest field strength then the anode current passes from T_2 to T_3 and A_{III} is connected with *D* by means of Re_3 .

It might be feared that the brief interruptions (a fraction of a second) during the relay action and the inevitable small variations in the sound volume would be troublesome, but in our practical experiments it appeared that this was hardly perceptible at all.

³⁾ A similar circuit but for only two valves has been given by Eccles and Jordan, *Radio Review*, **1**, 143, 1919; see also *Electronics* **12**, 14, Aug. 1939.

Counteracting interferences from other transmitters

Nature of the interferences

Another source of disturbances often making reception very unsatisfactory lies in the transmitting stations working on frequencies in the neighbourhood of the frequency of the station wanted. These interferences can be distinguished as 1) a whistling or, at times, a humming note, 2) the so-called cross-talk, and 3) side-band interference.

Whistling arises when the carrier wave signal from the interfering transmitter is insufficiently suppressed by the tuning circuits of the receiver; the frequency of that note equals the difference in frequencies of the station required and that causing the disturbance. Cross-talk occurs, for instance, when the carrier wave and the modulation of the disturbing station reach the detector in strength; the speech or the music from that station is then heard coming through the programme to which one is tuned in. Finally by side-band interference is understood the interference with a frequency lying between the frequency of the disturbing side-band and the carrier-wave frequency of the desired station; this results in an unpleasant, unintelligible, hissing noise, called "side-band splash" or "monkey-chatter".

Intervals of 9000 c/s are prescribed for the carrier wave frequencies of broadcasting stations on the medium and long wave ranges. For the sake of sound quality many stations go much higher than half this interval with their modulation frequency (e.g. to 10 000 c/s), and consequently at places where their field strength is high they are apt to cause side-band interference in receiving sets tuned in to a station on an adjacent wavelength.

In order to reduce these disturbances one often has recourse to a diminution of the bandwidth passed through, in the high and medium frequency parts of the receiver with the aid of tuning circuits or band-filters, and in the low frequency part with a tone-filter, both of these means being adjustable within certain limits. Obviously this implies that the high notes in the programme from the desired station are suppressed or attenuated.

In good radio sets the sound reproduction is uniform up to frequencies of max. 5 000-6 000 c/s (a compromise between quality of reproduction and the number of broadcasting stations working simultaneously in a certain range). Most listeners, however, adjust by ear their bandwidth and tone controls in such a way that frequencies above say 1500 c/s are already considerably attenuated. The music then sounds muffled and speech is not clear,

so that they then try to remedy this by increasing the volume, sometimes to the annoyance of others.

The less the frequency difference between the station desired and the interfering station, the more serious these troubles become.

Receiving antenna with directional effect

Interferences arising from other stations than the one desired can be reduced much more effectively — and, moreover, without sacrificing quality of sound — by employing directed reception. For this purpose antennae or combinations of antennae are used which have "directional effect", that is to say those which are more sensitive to signals from certain directions than to those from other directions. The antenna system is mounted or adjusted in such a way that its "sensitivity" is great for signals coming from the direction of the station desired and small for those coming from the direction of the interfering station. This method only fails in the rare case where both stations lie in the same direction. Preferably one would like to use an aerial system that receives signals from one direction only, or at least from directions at small angles from each other in the horizontal plane. Such antennae with a "sharp directional effect" are possible, but their dimensions are greater than the wavelength to be received, so that generally they can only be used on the ultra-short wave range. For the ordinary broadcasting wavelength (200-2000 m) the dimensions and cost would be prohibitive, the more so if it should be necessary to turn the aerial system in the directions of the various transmitting stations.

For the broadcasting range in question, therefore, antenna systems for directed reception must be used which are of small dimensions compared with the wavelength. This means that one has to be satisfied with antenna systems having such a directional effect that little or nothing is picked up of transmissions coming from one or two directions (according to the system employed), and for that direction (or directions) one chooses that of the interfering station(s). Such systems are: a) the frame aerial, and b) the frame aerial in combination with an ordinary antenna. We will now consider these two methods of directional reception more closely.

a) The single frame aerial

The magnetic field strength of an electromagnetic wave (of not too small wavelength) propagated along the earth's surface is directed horizontally and perpendicular to the direction of propagation of the wave. A frame aerial responds only to that

component of the magnetic field that is perpendicular to the plane of the frame. This component, therefore, disappears when that plane is perpendicular to the direction of the transmitting station. If the frame is rotated around a vertical axis (other axes are not to be considered at all here), then we get for the amplitude V_a of the voltage excited by the magnetic field the horizontal directional diagram represented in *fig. 3*, which can be expressed by the equation

$$V_a = V_0 \cos \alpha \dots \dots (1)$$

in which α is the angle of the plane of the frame aerial to the direction of the transmitter; if it is turned in that direction ($\alpha = 0$) one gets the maximum amplitude V_0 .

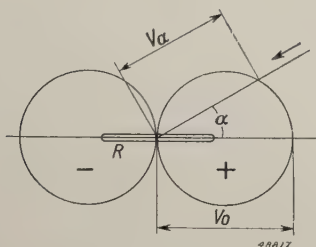


Fig. 3. Horizontal directional diagram of a frame aerial (R). V_a denotes the voltage excited in the aerial by a transmitter whose direction is at an angle α with the plane of the frame. This voltage becomes V_0 when $\alpha = 0$. The circle marked $+$ is for waves coming from the right, that marked $-$ for those coming from the left.

From *fig. 3* it is to be seen that when turning the frame 360° there are two sharp minima in the strength of reception. It is to be supposed, therefore, that by employing a frame aerial one could eliminate entirely the interferences from an outside station, provided the direction of that station makes an angle with the direction of the station desired which does not differ too much from 90° . This, however, is only correct if steps have been taken to attenuate sufficiently the so-called antenna effect of the frame ⁴⁾; some of these measures will be mentioned later on.

By antenna effect is understood the phenomenon that unless the said measures are applied the frame aerial may function simultaneously also as an ordinary or capacitive antenna. That is to say, also the electric component of an electromagnetic wave may excite a voltage in the frame regardless of the direction of the frame. In the diagram this finds expression, *i.e.* in the fact that the minima become less deep (thus no longer nil), so that a certain amount of signal of an interfering station always remains. In the articles referred to in footnote ⁴⁾ it is indicated how this antenna effect is best avoided.

⁴⁾ M. Ziegler, *Philips Techn. R.* **2**, 216, 1937, and P. Cornelius, *Philips Techn. R.* **7**, 65, 1942.

Here we will briefly mention what is referred to as night effect (so called because it coincides with fading effect, which occurs mostly at night). This manifests itself in the phenomenon that in moments of fading the signals from a transmitter sometimes appear to come from directions deviating more or less from the actual direction of the station. Even if the reception has been cleared of disturbing signals by turning the frame aerial the right way in the absence of fading, those signals may come through again as soon as fading effect occurs. Considering the great irregularity in the occurrence of fading effect there is little sense in trying to correct the position of the frame. Taken on an average, the disturbance-free position, determined while fading is absent, is best.

b) Combination of a frame aerial and an ordinary antenna; cardioid and cycloidal reception

When the station desired and the interfering station lie in approximately opposite directions with respect to the receiving station, a frame aerial alone will not suffice, but now a combination of a frame aerial and an ordinary antenna can be used, as will be explained in the following.

Let us suppose that a transmitter may again excite a voltage with amplitude V_0 in the frame aerial turned in its direction. A circuit connected to an ordinary antenna likewise installed is adjusted in such a way that the transmitter excites a voltage V_0 also in that antenna. The two voltages can be added together so as to give an amplitude $2V_0$. Upon the frame then being turned 180° the phase of the voltage is reversed, so that the frame aerial voltage and the antenna voltage neutralise each other and no signals of the transmitter in question remain. Reception from other directions is still possible; the equation for the combined voltage V_{res} is:

$$V_{\text{res}} = V_0 (1 + \cos \alpha) \dots \dots (2)$$

This equation is represented by the heart-shaped curve of *fig. 4*, from which the term "cardioid reception" is derived.

In the foregoing it has been tacitly supposed that the voltage of the frame aerial and that of the ordinary antenna differ 0° or 180° in phase. Actually, however, there is between these two a phase difference of $+$ or -90° . It is therefore necessary to introduce a device that changes the phase of the antenna voltage for instance 90° . This device can be combined with an amplitude control by means of which the antenna voltage output can be adjusted to the right value. An

example of how this can be done will be given farther on.

As may be seen from fig. 4, the directional diagram for the method described has one zero direction. By a slight modification, however, it is possible to get two such directions making any

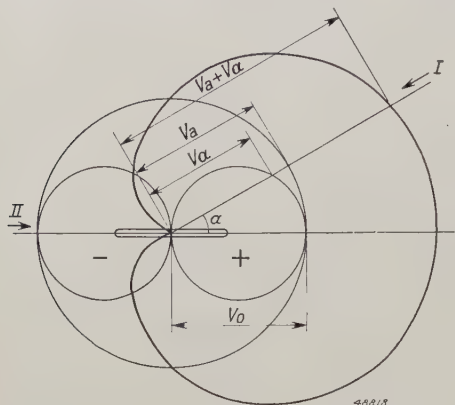


Fig. 4. Directional diagram for a combination of a frame aerial R (circles $+$ and $-$ with diameter V_0 , cf. fig. 3) and an ordinary antenna without directional effect (large circle, radius V_a). If V_a is made equal to V_0 and steps are taken to ensure that V_a and V_0 are in phase or in counter-phase as the case may be, then we get the heart-shaped diagram (cardioid reception). A signal $V_a + V_0 = V_0(1 + \cos \alpha)$ is obtained from direction I ; nothing is received from direction II .

desired angle, simply by making the amplitude of the antenna voltage V_a smaller than that of the maximum voltage V_0 in the frame aerial. The directional diagram then obtained is an extended epicycloid, and for that reason this system is referred to as cycloidal reception (the cardioid is a special case of the epicycloid). The directions from which nothing comes through, $\pm \alpha_{\min}$, then follow from

$$V_a + V_0 \cos \alpha_{\min} = 0. \dots (3)$$

By turning the frame so that its plane bisects the angle between the directions of two interfering stations, it is in this way theoretically possible to suppress the interferences from both of them.

In practice it is not quite so simple to suppress two interfering signals. As already stated, one must have a device for deriving from the antenna voltage a voltage shifted 90° . The phase displacement obtained, however, is more or less dependent upon the frequency, and as a rule the two interfering stations will not be working on the same frequency. We cannot here enter upon the complications arising if one nevertheless tries to make the phase displacement 90° for each of the stations, as is necessary for reception free from interference.

Reverting to the case of one interfering station, we may add that according to equation (3) the interfering signals can be reduced to zero also for any arbitrary position α of the frame, by making

$V_a/V_0 = -\cos \alpha$. Supposing that this condition is always satisfied, it can be ascertained what value of α yields the strongest signal of the station desired. For that purpose the plane of the frame must be perpendicular to the line bisecting the angle between the directions of the two transmitters, as demonstrated by the following calculation.

In fig. 5 φ denotes the angle between the directions of the desired station (I) and the interfering station (II). The frame is set in an arbitrary position where its plane forms an angle α with the direction II . It is required to ascertain the optimum value of α , that is to say the value which combines complete suppression of the interferences from II with the strongest signal from I .

If we use V_a'' for the amplitude of the voltage excited by the interfering station II in the frame aerial in the position given, and V_0'' for that amplitude when the frame was directed to II , then according to equation (1) we get

$$V_a'' = V_0'' \cos \alpha.$$

The ordinary antenna yields a voltage V_a'' derived from II and adjustable in size and phase. We assume that V_a'' is always adjusted in counter-phase to V_0'' . The condition necessary for eliminating interferences from II is then

$$V_a'' = V_0'' = V_0'' \cos \alpha \dots (4)$$

Now in order to find the value of σ which not only satisfies (4) but also yields the strongest signal

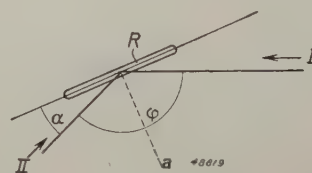


Fig. 5. The directions of the desired transmitter (I) and the interfering transmitter (II) make an angle φ ; the plane of the frame aerial forms an angle α with the direction II . The optimum reception, with a combination of the frame aerial and an ordinary antenna, is obtained when the plane of the frame is perpendicular to the bisectrix of the angle φ ($\alpha = 90^\circ - \frac{1}{2} \varphi$).

from the station I we make use of the fact that in an electromagnetic wave — i.e. in a field of irradiation at a sufficiently great distance (e.g. more than 10 wavelengths) from the transmitting antenna — there is a fixed relation between the electric and the magnetic field strength⁵⁾. Since

⁵⁾ In a travelling wave the ratio of the electric to the magnetic field strength is equal to $120 \pi \approx 377$, if Giorgi's rationalised units are used with absolute volts and amperes, and the said field strengths are therefore expressed in V/m and A/m respectively.

the ordinary antenna responds to the electric component of the wave and the frame aerial to the magnetic component, there is therefore the same relation between the voltages set up by different transmitters in the ordinary antenna and maximally in the frame aerial. By introducing for the voltages derived from the desired station *I* the symbols V_0' , V_a' and $V_{a'}$, with the same significations as previously given, then we get:

$$\frac{V_{a'}}{V_0'} = \frac{V_{a''}}{V_0''} \dots \dots \dots (5)$$

Further, having regard to (1)

$$V_{a'} = V_0' \cos \{180^\circ - (\alpha + \varphi)\} = -V_0' \cos (\alpha + \varphi).$$

Making use of (5) we may write for (4):

$$V_{a'} = V_0' \cos \alpha.$$

The resultant voltage of the desired station *I* is therefore:

$$\begin{aligned} V_{\text{res}}' &= V_{a'} + V_{a'} = V_0' \{-\cos (\alpha + \varphi) + \cos \alpha\} = \\ &= 2 V_0' \sin \frac{1}{2} \varphi \sin (\alpha + \frac{1}{2} \varphi). \end{aligned}$$

This expression reaches its maximum value for $\alpha = 90^\circ - \frac{1}{2} \varphi$, viz:

$$V_{\text{res max}}' = 2 V_0' \sin \frac{1}{2} \varphi \dots \dots (6)$$

The condition $\alpha = 90^\circ - \frac{1}{2} \varphi$ can also be expressed as follows: in order to get the maximum signal strength from the station *I* without receiving any disturbing signals from station *II*, using cycloidal reception, the frame has to be turned in such a way that it takes a position perpendicular to the line (α in fig. 5) bisecting the angle between the directions of *I* and *II*.

On the other hand with cardioid reception (frame turned in the direction of the interfering station *II*: $\alpha = 0$) the magnitude of the resulting voltage from the desired station *I* would amount to

$$V_{\text{res card.}}' = V_0' (1 - \cos \varphi) = 2 V_0' \sin^2 \frac{1}{2} \varphi, \quad (7)$$

which is smaller by a factor $\sin \frac{1}{2} \varphi$ compared with the value (6). Therefore the two methods are only equivalent when $\varphi = 180^\circ$.

For the sake of completeness we will now consider the case where exclusively a frame aerial is used. To make the reception free of interferences α must be made equal to 90° (fig. 5), so that the plane of the frame forms an angle $\varphi - 90^\circ$ with the direction of the desired station *I*. The latter thus excites in the frame aerial a voltage $V_{a'}$, given by:

$$\begin{aligned} V_{a'} &= V_0' \cos (\varphi - 90^\circ) = \\ &= V_0' \sin \varphi = 2 V_0' \cos \frac{1}{2} \varphi \sin \frac{1}{2} \varphi. \quad (8) \end{aligned}$$

In fig. 6 we have plotted the results of (6), (7) and (8) — omitting the factor $2 V_0'$ occurring in all three — as functions of the angle φ between the directions of the two stations. As will be seen, the cardioid reception is better than that with the

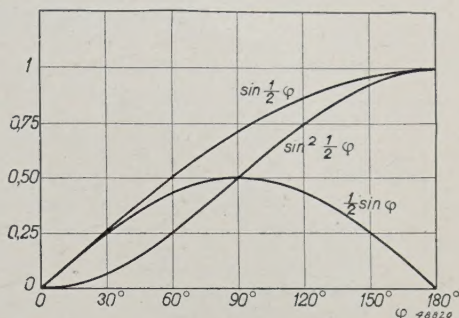


Fig. 6. The resulting signal of the desired station (omitting the factor $2 V_0'$) is plotted as a function of the angle φ between the directions of the desired station and the interfering stations, for three methods of reception. The disturbing signal is in every case entirely suppressed. The curve $\frac{1}{2} \sin \varphi$ ($= \sin \frac{1}{2} \varphi \cos \frac{1}{2} \varphi$) applies for a frame aerial alone, the curve $\sin^2 \frac{1}{2} \varphi$ for cardioid reception, the curve $\sin \frac{1}{2} \varphi$ for cycloidal reception, where, among others, the condition $\alpha = 90^\circ - \frac{1}{2} \varphi$ (cf. fig. 5) is satisfied.

frame aerial alone when the difference in direction φ is more than 90° ; if it is smaller the reverse is the case. The cycloidal method, where the frame is perpendicular to the bisectrix of φ ($\alpha = 90^\circ - \frac{1}{2} \varphi$), is in every case better than the other two methods. From fig. 6 it is to be concluded that where there is only a small directional difference φ , there is little advantage in the cycloidal method compared with what can be obtained with a single frame aerial. It is not to be forgotten, however, that with the latter method any small antenna effect remaining will make it impossible to suppress the interferences entirely. Such is indeed possible with the former method, as also with cardioid reception, which, however, according to fig. 6 gives a considerably smaller signal strength when φ is small.

Description of an apparatus for improved broadcasting reception

The receiving set and the aerial couplings

Philips have designed a special receiving set for the methods of reception described (type 4578, fig. 7). This set is characterised by a highly uniform amplification in a bandwidth adjustable in four stages and extending to a maximum of $2 \times 10\,000$ c/s (thus much farther than is the case with normal broadcasting receivers, where it would be of little use owing to the interferences).

There are some variations in design of this model; design 4578/03 is specially constructed for cycloidal and cardioid reception. Fig. 8 shows how the first

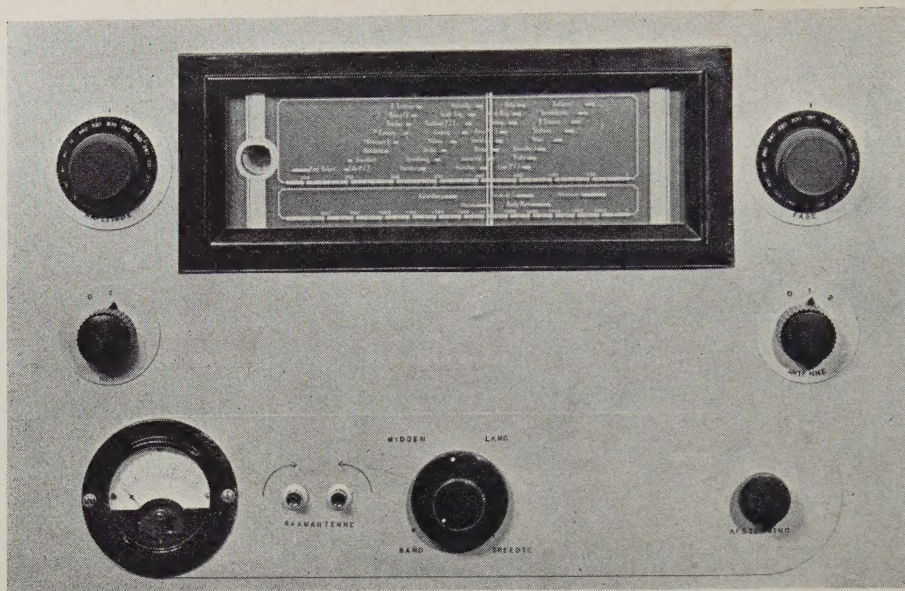


Fig. 7. Receiving set (type 4578/03) for cycloidal and cardioid reception. On either side of the station dial is a knob with which the signal from the ordinary antenna is given the right amplitude and phase with respect to the signal from the frame aerial. Below these knobs there is on the left the mains switch and on the right a knob which in the position "0" switches off the ordinary aerial and in the positions "1" and "2" switches it on; when changing over from "1" to "2" the aerial coupling coil is commuted, so that the frame aerial need never be turned more than 180° . At the bottom, from left to right: a meter indicating the position of the frame (scale $0-180^\circ$); two buttons for the motor turning the frame (two directions of rotation); a double knob, the outer one operating the switch for "medium wave" and "long wave" and the inner one regulating the bandwidth in four stages (to max. $2 \times 10\,000$ c/s); on the extreme right the tuning knob.

h.f. circuit is connected with the ordinary antenna and with the frame aerial; it is described in further detail in the text underneath the diagram.

The frame aerial and its connection to the set

In order to reduce the "antenna effect" of the frame aerial (see note ⁴) two measures have been devised: the frame (see fig. 9) consists of only one winding, and the primary coil of the transformer T (fig. 8) is earthed in its middle. This transformer is necessary to step up the voltage excited in the frame aerial and thus adapt the self-inductance of the frame (which is relatively very low with a single winding) to the capacity of the normal tuning condenser. The coupling of the coils of this transformer has to be very close. For that reason a ferromagnetic core is used; "Ferroxcube" ⁶⁾ is a very suitable material for this.

As to the setting up of the frame, it should preferably be located at some distance from the receiving set, because of the local interferences often transmitted by the mains to which the set is connected ⁷⁾. It is desirable, however, that the condenser tuning the frame circuit should be left in the set for the sake of single-knob tuning. This

means that a transmission line has to connect the set with the frame aerial.

This transmission line should have an exceptionally low resistance, because owing to the very close transformer coupling this resistance, multiplied by the square of the transformer ratio, appears in the tuning circuit as a loss resistance. In the

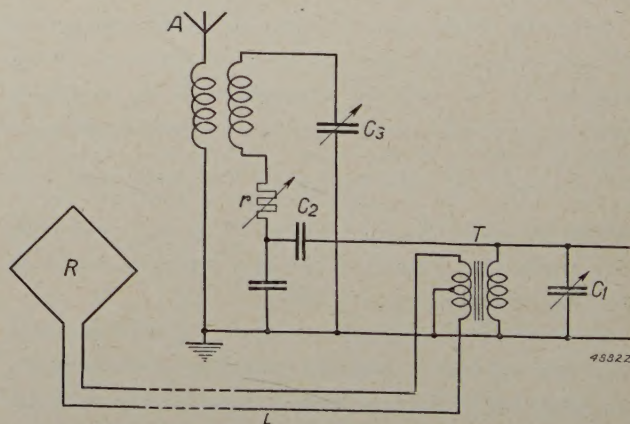


Fig. 8. The frame aerial R is connected *via* a transmission line L of 12 m length to a transformer T and to the variable condenser C_1 and forms the first tuned circuit of the receiving set. Across the small coupling condenser C_2 the signals from the ordinary antenna A are led in at the same time. These signals can be adjusted in amplitude by means of the resistance r and in phase by means of the condenser C_3 . The circuit of which C_3 forms a part has to be approximately tuned to the wavelength of the interfering station.

⁶⁾ Philips Techn. R. 8, 353, 1946.

⁷⁾ Philips Techn. R. 3, 235, 1938 and 6, 302, 1941.

second place, considering its divided self-inductance and capacity, together with the frame, it should behave approximately like a coil with constant

of single-knob tuning, it is desired to keep tuned together with the other h.f. circuits by means of a multiple variable condenser. This implies that the transformed self-inductance of the frame and of the line — bearing in mind the finite self-inductance and leakage of the transformer — must be equal to the self-inductance in the other h.f. circuits. In the third place the self-inductance of the line should be small compared with that of the aerial, so that the aerial voltage may be conducted to the transformer with the least loss.

These requirements are partly contradictory, but a satisfactory solution has been found, firstly by making the line of such a length that, while on the one hand the frame aerial is far enough removed from the mains interferences, on the other hand the longest characteristic wave of the line together with the frame aerial is sufficiently smaller than the shortest wave to be received (*i.e.* 200 m); and further by choosing the self-inductance of the line at such a value that it is only a fraction of the self-inductance of the frame aerial. The length chosen on these arguments was 12 meters. The transmission line itself consists of two parallel copper strips 25 mm wide, placed 5 mm apart and mainly separated by air. The losses prove to be small enough; the self-inductance is about one-third of that of the frame aerial. Compared with a concentric line this solution has the advantage that it is symmetrical with respect to the surroundings and consequently does not produce any antenna effect.

Diversity reception

In the beginning of 1940 an installation was set up at Rotterdam on the principles of fig. 1, but no frame was used. As regards reduction of selective fading effect the results were most satisfactory. Owing to the war the development of this method of reception was temporarily suspended. It is to be expected, however, that in the near future more interest will be shown in this method, the more so since now, thanks to the development of directional reception, the main sources of interferences can be successfully counteracted by a combination of the two systems.

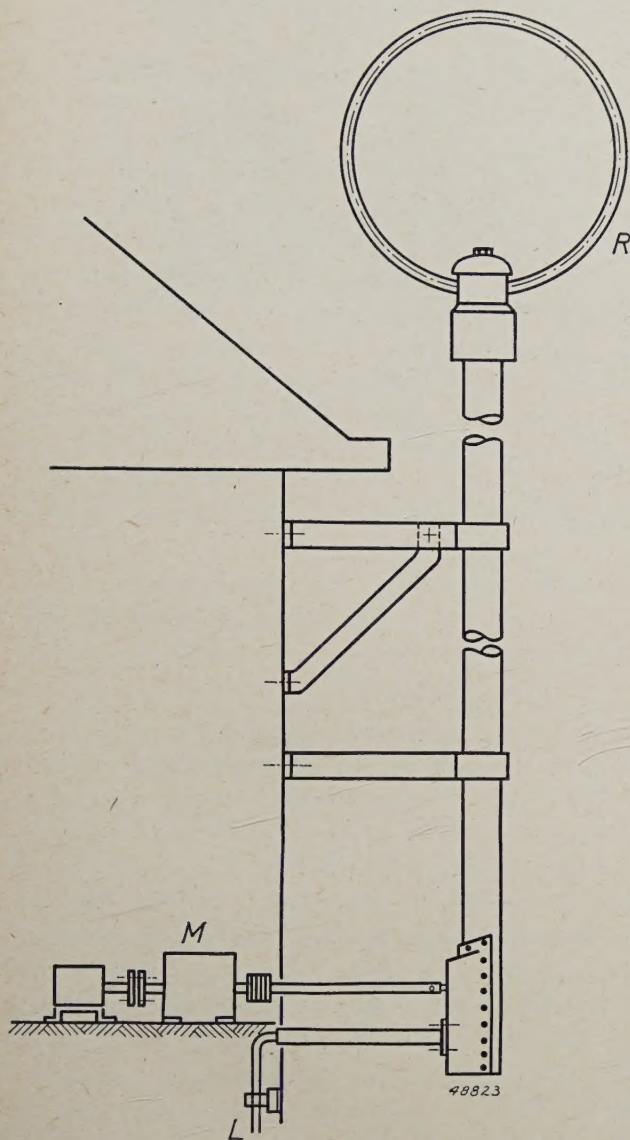


Fig. 9. Setting up the frame aerial on a building. *R* = frame aerial (one winding of copper tubing; diameter at least 1 m). *M* = motor turning the frame. *L* = transmission line connecting to the receiver.

self-inductance and constant self-capacity in the whole of the frequency range to be covered; it forms part of the first h.f. circuit, which, for the sake

ABSTRACTS OF RECENT SCIENTIFIC PUBLICATIONS OF THE N.V. PHILIPS' GLOEILAMPENFABRIEKEN

Reprints of the majority of these papers can be obtained on application to the Administration of the Research Laboratory, Kastanjelaan, Eindhoven, Netherlands. Those papers of which no reprints are available in sufficient number are marked with an asterisk.

- 1708:** P. C. van der Willigen: Contact arc welding. (The Welding Journal Research Supplement, May 1946).

This paper describes the properties and the advantages of a new type of welding electrode (developed by Philips), which combines the properties of touch welding, self-starting and reignition while also some other advantages are obtained (for full particulars see Philips Technical Review **8**, 161-167, 304-309, 1946).

- 1709:** N. G. de Bruyn: On the Zeros of a polynomial and its derivative (Proc. Kon. Ned. Akad. Wet. Amsterdam **49**, 1037-1044 1946).

It is proved that the sum of the absolute values of the imaginary parts of the roots of $f'(z)$ is equal to or less than $(1-1/n)$ times the corresponding sum for $f(z)$, in which $f(z)$ is a real polynomial of degree n in z and $f'(z)$ its derivative. This theorem is extended and a few specializations are considered. It is unknown whether the inequality holds for polynomials with complex coefficients. It does if all the roots are assumed to lie on the imaginary axis.

- 1710:** H. B. G. Casimir and D. Polder: Influence of retardation on the London-van der Waals forces (Nature **158**, 787, 1946).

In the course of work on the stability of colloidal solutions, Overbeek arrived at the conclusion that

in order to obtain agreement between theory and experiment it is necessary to assume that the London-van der Waals energy decreases more rapidly than R^{-6} . He pointed out that the retardation of the electrostatic forced might be responsible for such an effect and that deviations from the R^{-6} law should become effective at a distance comparable to the wavelength λ corresponding to the excitation energies of the interacting atoms. Calculations with the aid of quantum electrodynamics prove that this suggestion holds true, the energy being proportional to R^{-7} rather than to R^{-6} for $R \gg \lambda$. Details of the quantum-mechanical calculation will be published in the Physical Review.

- 1711:** P. J. Bouma: Die Grassmannschen Gesetze der Farbmischung (Physica **12**, 545-552, 1946) (Grassmann's laws of additive colour mixing).

The author proposes a formulation of Grassmann's laws of additive colour mixing, which is axiomatically as pure as possible and which does not contain more than is necessary for the construction of elementary colorimetry. The correctness of the formulation is proved by working out this construction in rough lines finally the questions of continuity, arising in connexion herewith, are dealt with and are reduced to the existence of thresholds.

The last issue of *Philips Research Reports* (Nr 6 of volume 1, December 1946) contains the following papers:

- R27: N. G. de Bruyn: A note on van der Pol's equation.
 R28: T. Jurriaanse: The influence of gas density and temperature on the normal cathode fall of a gas discharge in rare gases.
 R29: J. W. L. Köhler and C. G. Koops: A new alternating current bridge for precision measurements.
 R30: M. Gevers: The relation between the power factor and the temperature coefficient of the dielectric constant of solid dielectrics, V.
 R31: K. F. Niessen: A new method for the determination of the electric constants of the earth's surface.

Readers interested in one of the above mentioned articles may apply to the Administration of the Philips Physical Laboratory, Kastanjelaan Eindhoven, where a limited number of copies are available for distribution. For a subscription to Philips Research Reports please write to the publishers of Philips Technical Review.